

Case 7170

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of : March 7, 2003  
STEFFEN SETZER et al. :  
Serial No.: 09/926,225 :  
I.A. No.: PCT/EP00/02639 :  
I.A. Filing Date: 27 March 2000 :  
For: DEVICE FOR PICTORIALY :  
DEPICTING THREE-DIMENSIONAL :  
OBJECTS :

RECEIVED  
18 MAR 2003  
Legal Coun.  
International Division

PETITION TO WITHDRAW NOTICE OF ABANDONMENT

Honorable Commissioner of Patents  
and Trademarks  
Washington, D.C. 20231

Dear Sir:

Applicants received a Notice of Abandonment mailed on February 25, 2003, copy attached, in the above-identified application, "for failure to timely or properly reply to the Notice to File Missing Parts (Notice) mailed on 11/08/2001." Applicants believe that the Notice of Abandonment was sent in error.

Applicants filed a response to the Notice on December 28, 2001, well within the due date of January 8, 2002 for filing a response to the Notice. Attached herewith are copies of our response, including a copy of our receipt card indicating that a Response was filed on December 28, 2001. Also attached is a copy of the canceled check for the surcharge required by the Notice.

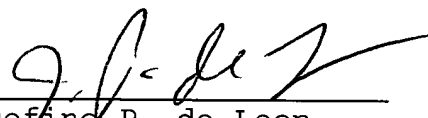
Serial No. 09/926,225

- 2 -

In view of the above, applicants hereby request that the Notice of Abandonment be withdrawn.

It is believed that no fee is due; however, should that be incorrect, please charge Deposit Account No. 19-2105 and inform the undersigned.

Respectfully submitted,

  
\_\_\_\_\_  
Josefino P. de Leon  
Reg. No. 33,166

SHLESINGER, ARKWRIGHT & GARVEY LLP  
3000 South Eads Street  
Arlington, Virginia 22202  
(703) 684-5600  
lm



UNITED STATES  
PATENT AND  
TRADEMARK OFFICE

Commissioner for Patents  
Washington, DC 20231  
www.uspto.gov

| APPLICATION NUMBER | FILING DATE | FIRST NAMED APPLICANT | ATTY. DOCKET NO./TITLE |
|--------------------|-------------|-----------------------|------------------------|
| 09/926,225         |             | Steffen Setzer        | 7170                   |

Josefino P. De Leon  
Shlesinger Arkwright & Garvey  
3000 South Eads Street  
Arlington, VA 22202

CONFIRMATION NO. 2162

**ABANDONMENT/TERMINATION  
LETTER**



\*OC000000009558744\*

Date Mailed: 02/25/2003

**NOTICE OF ABANDONMENT UNDER 37 CFR 1.53 (f) OR (g)**

The above-identified application is abandoned for failure to timely or properly reply to the Notice to File Missing Parts (Notice) mailed on 11/08/2001.

- No reply was received.

A petition to the Commissioner under 37 CFR 1.137 may be filed requesting that the application be revived.

Under 37 CFR 1.137(a), a petition requesting the application be revived on the grounds of **UNAVOIDABLE DELAY** must be filed promptly after the applicant becomes aware of the abandonment and such petition must be accompanied by: (1) an adequate showing of the cause of unavoidable delay; (2) the required reply to the above-identified Notice; (3) the petition fee set forth in 37 CFR 1.17(l); and (4) a terminal disclaimer if required by 37 CFR 1.137(d).

Under 37 CFR 1.137(b), a petition requesting the application be revived on the grounds of **UNINTENTIONAL DELAY** must be filed promptly after applicant becomes aware of the abandonment and such petition must be accompanied by: (1) a statement that the entire delay was unintentional; (2) the required reply to the above-identified Notice; (3) the petition fee set forth in 37 CFR 1.17(m); and (4) a terminal disclaimer if required by 37 CFR 1.137(d).

Any questions concerning petitions to revive should be directed to "Office of Petitions" at (703) 305-9282.

---

*A copy of this notice **MUST** be returned with the reply.*

---

Customer Service Center  
Initial Patent Examination Division (703) 308-1202

PART 2 - COPY TO BE RETURNED WITH RESPONSE

Case 7170

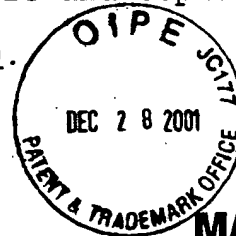
JPD/lm

In re Patent Application of STEFFEN SETZER et al.  
Serial No.: 09/926,225  
International Application No.: PCT/EP00/02639  
International Filing Date: 27 March 2000  
For: DEVICE FOR PICTORIALY DEPICTING THREE-DIMENSIONAL OBJECTS

INCLUDING: **RESPONSE** (2 pgs.)

Copy of the Notification of Missing Requirements  
Under 35 U.S.C. 371 in the United States  
Designated/Elected Office (DO/EO/US) (2 pgs.)  
Executed Declaration and Power of Attorney (3 pgs.)  
Information Disclosure Statement (1 pg.)  
Form PTO-1449 (1 pg.) - including copy of the  
International Search Report and copies of the  
cited four(4) references  
Check in the amount of \$65.00.

PLEASE STAMP DATE RECEIVED AND RETURN.



**MAILROOM**



**SHLESINGER ARKWRIGHT AND GARVEY, LLP**

3000 S EADS ST  
ARLINGTON, VA 22202

FIRST VIRGINIA BANK  
FALLS CHURCH, VIRGINIA 22042  
68-111-560

21804

12/28/2001

PAY TO THE  
ORDER OF Commissioner of Patents & Trademarks

\$ 65.00

Sixty Five and 00/100

DOLLARS

Commissioner of Patents &  
Trademarks

2900 Crystal Drive  
Arlington, VA 22202-4351

MEMO

Case 71-70

000021804 05600118 0704 3414

00000005500

ENDORSE HERE

PATENT AND TRADEMARK OFFICE

13-10-0001  
01-03-2002  
FOR CREDIT TO THE  
U.S. TREASURY

DO NOT WRITE, STAMP OR SIGN BELOW THIS LINE  
RESERVED FOR FINANCIAL INSTITUTION USE

0025 27511

FIRST VIRGINIA BANK  
FALLS CHURCH, VA  
68-111-560  
056001118

JUN - 4 02

PAID



The security features listed below are those  
not listed, exceed industry guide

RS-9

**Security Features:**

- MicroPrint Signature Line
- Results of document verification:
  - Small type in signature appears as dotted line when copied
  - Stains or spots appear with chemical alteration
  - White mark appears when erased
- Erasure Protection
- Security Screen
- Absence of "Original Document"

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of : December 28, 2001  
STEFFEN SETZER et al. :  
Serial No.: 09/926,225 :  
I.A. No.: PCT/EP00/02639 :  
I.A. Filing Date: 27 March 2000 :  
For: DEVICE FOR PICTORIALY :  
DEPICTING THREE-DIMENSIONAL :  
OBJECTS :

R E S P O N S E

Honorable Commissioner of Patents  
and Trademarks  
Washington, D.C. 20231

Dear Sir:

In response to the Notification of Missing Requirements Under 35 U.S.C. 371 in the United States Designated/Elected Office (DO/EO/US) of November 8, 2001, enclosed herewith are the following.

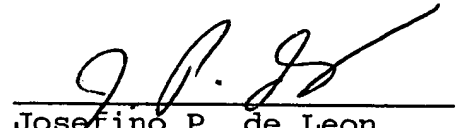
1. Declaration and Power of Attorney;
2. Copy of the Notification of Missing Requirements Under 35 U.S.C. 371 in the United States Designated/Elected Office (DO/EO/US); and
3. Check in the amount of \$65.00 for the surcharge for later filing of the declaration.

Serial No. 09/926,225

- 2 -

It is believed that no additional fee is due; however, should that be incorrect, please charge Deposit Account No. 19-2105 and inform the undersigned.

Respectfully submitted,



---

Josefino P. de Leon  
Reg. No. 33,166

SHLESINGER, ARKWRIGHT & GARVEY LLP  
3000 South Eads Street  
Arlington, Virginia 22202  
(703) 684-5600  
lm

DECLARATION AND POWER OF ATTORNEY - ORIGINAL APPLICATION

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below beneath my name:

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which Letters Patent is sought on the invention entitled

DEVICE FOR PICTORIALY DEPICTING THREE-DIMENSIONAL OBJECTS

the specification of which \_\_\_\_\_ is attached hereto  
(check one) X was filed on 27 March 2000  
as Application S.N. PCT/EP00/02639  
and was amended on \_\_\_\_\_

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by an amendment referred to above.

I acknowledge the duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56(a).

I hereby claim foreign priority benefits under Title 35, United States Code, §119(a)-(d) or §365(b) of any foreign application(s) for patent or inventor's certificate, or §365(a) of any PCT International Application which designated at least one country other than the United States, listed below and have also identified below any foreign application for patent or invention certificate, or a PCT International application having a filing date before that of the application on which priority is claimed.

Prior Foreign Application(s)                      Priority Claimed

| (Number) | (Country) | (Day/Month/Year Filed) | YES | NO |
|----------|-----------|------------------------|-----|----|
|----------|-----------|------------------------|-----|----|

|          |           |                        |     |    |
|----------|-----------|------------------------|-----|----|
| (Number) | (Country) | (Day/Month/Year Filed) | YES | NO |
|----------|-----------|------------------------|-----|----|

I hereby claim the benefit under 35 U.S.C. §119(e) of any United States provisional application(s) listed below:

---

(Application Ser. No.)

(Filing Date)

---

(Application Ser. No.)

(Filing Date)

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s), or §365(c) of any PCT International application designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application or PCT International application in the manner provided by the first paragraph of Title 35, United States Code, §112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, §1.56(a) which occurred between the filing date of the prior application and the national or PCT international filing date of this application.

---

(Application Ser. No.)

(Filing Date)

(Status)

(patented, pending, abandoned)

POWER OF ATTORNEY: As a named inventor, I hereby appoint the following attorney(s) and/or agent(s) to prosecute this application and transact all business in the Patent and Trademark Office connected therewith:

B. Edward Shlesinger, Jr., Reg. No. 17,225; George A. Arkwright, Reg. No. 20,158; George A. Garvey, Reg. No. 17,737; Josefino P. de Leon, Reg. No. 33,166; Terrence L. B. Brown, Reg. No. 32,685; Michael M. Zadrozny, Reg. No. 30,985; and Brian J. Marton, Reg. Agent 30,292;

SEND CORRESPONDENCE TO: SHLESINGER, ARKWRIGHT & GARVEY LLP  
3000 South Eads Street  
Arlington, Virginia 22202

DIRECT TELEPHONE CALLS TO Josefino P. de Leon (703) 684-5600

The undersigned hereby authorizes the U.S. attorney(s) or agent(s) named herein to accept and follow instructions from Hössle & Kudlek as to any action to be taken in the Patent and Trademark Office regarding this application without direct communication between the U.S. attorney(s) or agent(s) and the undersigned. In the event of a change in the persons from whom instructions may be taken, the U.S. attorney(s) or agent(s) named herein will be so notified by the undersigned.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that the statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under §1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

Full name of sole or first inventor Steffen SETZER

Inventor's signature X 22.11.07 Steffen Setzer Date 22.11.07

Residence Am Wallgraben 142, D-70565 Stuttgart, Germany

Citizenship German

Post Office Address same as above

Full name of second or joint inventor Thilo WEIGEL

Inventor's signature X 13.12.2001 Thilo Weigel Date 13.12.2001

Residence Am Wallgraben 142, D-70565 Stuttgart, Germany

Citizenship German

Post Office Address same as above

Full name of third or joint inventor Dittmar WRIGHT

Inventor's signature X Dittmar Wright Date 2.3.11.01

Residence Am Wallgraben 142, D-70565 Stuttgart, Germany

Citizenship German

Post Office Address same as above

|   |  |  |
|---|--|--|
| <input type="checkbox"/> <small>Check for Patent, See PCT, United States Patent and Trademark Office, Washington, D.C. 20590, www.uspto.gov</small> |  |  |
|---|--|--|

|                             |                       |                  |
|-----------------------------|-----------------------|------------------|
| U.S. APPLICATION NUMBER NO. | FIRST NAMED APPLICANT | ATTY. DOCKET NO. |
|-----------------------------|-----------------------|------------------|

09/926,225

Steffen Setzer

7170

|                               |
|-------------------------------|
| INTERNATIONAL APPLICATION NO. |
|-------------------------------|

PCT/EP00/02639

|                  |               |
|------------------|---------------|
| I.A. FILING DATE | PRIORITY DATE |
|------------------|---------------|

03/27/2000

Josefino P. De Leon  
Shlesinger Arkwright & Garvey  
3000 South Eads Street  
Arlington, VA 22202

RECEIVED

NOV 13 2001

SHLESINGER, ARKWRIGHT  
& GARVEY LLP

CONFIRMATION NO. 2162

371 FORMALITIES LETTER



\*OC000000007035841\*

Date Mailed: 11/08/2001

### NOTIFICATION OF MISSING REQUIREMENTS UNDER 35 U.S.C. 371 IN THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US)

The following items have been submitted by the applicant or the IB to the United States Patent and Trademark Office as an Elected Office (37 CFR 1.495):

- U.S. Basic National Fees
- Indication of Small Entity Status
- Copy of IPE Report
- Copy of references cited in ISR
- Copy of the International Application
- Copy of the International Search Report
- Preliminary Amendments

The following items **MUST** be furnished within the period set forth below in order to complete the requirements for acceptance under 35 U.S.C. 371:

- Oath or declaration of the inventors, in compliance with 37 CFR 1.497(a) and (b), identifying the application by the International application number and international filing date.
- \$65 Surcharge for providing the oath or declaration later than the appropriate 30 months from the priority date (37 CFR 1.492(e)) is required.

ALL OF THE ITEMS SET FORTH ABOVE MUST BE SUBMITTED WITHIN TWO (2) MONTH FROM THE DATE OF THIS NOTICE OR BY 22 or 32 MONTHS (where 37 CFR 1.495 applies) FROM THE PRIORITY DATE FOR THE APPLICATION, WHICHEVER IS LATER. FAILURE TO PROPERLY RESPOND WILL RESULT IN ABANDONMENT.

The time period set above may be extended by filing a petition and fee for extension of time under the provisions of 37 CFR 1.136(a).

DKZ  
RSM

**SUMMARY OF FEES DUE:**

Total additional fees required for this application is \$65 for a Small Entity:

- \$65 Late oath or declaration Surcharge.

Applicant is reminded that any communications to the United States Patent and Trademark Office must be mailed to the address given in the heading and include the U.S. application no. shown above (37 CFR 1.5)

ANITA D JOHNSON

Telephone: (703) 305-3661

**PART 1 - ATTORNEY/APPLICANT COPY**

| U.S. APPLICATION NUMBER NO. | INTERNATIONAL APPLICATION NO. | ATTY. DOCKET NO. |
|-----------------------------|-------------------------------|------------------|
| 09/926,225                  | PCT/EP00/02639                | 7170             |



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of : December 28, 2001  
STEFFEN SETZER et al. :  
Serial No.: 09/926,225 :  
I.A. No.: PCT/EP00/02639 :  
I.A. Filing Date: 27 March 2000 :  
For: DEVICE FOR PICTORIALY :  
DEPICTING THREE-DIMENSIONAL :  
OBJECTS :

INFORMATION DISCLOSURE STATEMENT

Honorable Commissioner of Patents  
and Trademarks  
Washington, D.C. 20231

Dear Sir:

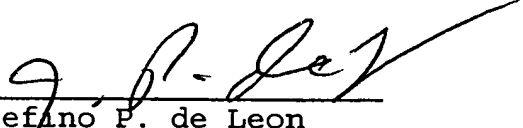
Attached herewith is a completed PTO Form-1449 and copies of the cited references. The Examiner is respectfully requested to consider these references and provide a signed copy of the form in the next communication in regard to the above-identified application. An International Search Report in the corresponding PCT application is attached herewith, showing the relevance of the references.

Serial No.: 09/926,225

-2-

It is believed that no fee is due; however, should that be incorrect, please charge Deposit Account No. 19-2105 and inform the undersigned.

Respectfully submitted,

  
\_\_\_\_\_  
Josefino P. de Leon  
Attorney for Applicant  
Reg. No. 33,166

SHLESINGER, ARKWRIGHT & GARVEY LLP  
3000 South Eads Street  
Arlington, Virginia 22202  
(703) 684-5600  
lm

|                            |                              |
|----------------------------|------------------------------|
| Form PTO-1449              | Docket No. : 7170            |
| INFORMATION DISCLOSURE     | Appl. No. : 09/926,225       |
| CITATION IN AN APPLICATION | Filing Date :                |
|                            | Applicant : S. SETZER et al. |
|                            | Group Art Unit:              |

## U.S. PATENT DOCUMENTS

| Ex.<br>Init. | Document<br>No. | Date       | Name        | Class/<br>Subclass | Filing<br>Date |
|--------------|-----------------|------------|-------------|--------------------|----------------|
|              | 5,818,959       | 10/06/1998 | Webb et al. |                    |                |
|              |                 |            |             |                    |                |
|              |                 |            |             |                    |                |
|              |                 |            |             |                    |                |
|              |                 |            |             |                    |                |
|              |                 |            |             |                    |                |
|              |                 |            |             |                    |                |
|              |                 |            |             |                    |                |
|              |                 |            |             |                    |                |
|              |                 |            |             |                    |                |

## FOREIGN PATENT DOCUMENTS

| Ex.<br>Init. | Document No. | Date     | Country | Class/<br>Subclass | Translation<br>Yes/No |  |
|--------------|--------------|----------|---------|--------------------|-----------------------|--|
|              | EP 0631250A2 | 28.12.94 | EPO     |                    |                       |  |
|              | WO 98/17970  | 30.04.98 | WIPO    |                    |                       |  |
|              | WO 98/09253  | 05.03.98 | WIPO    |                    |                       |  |

## OTHER DOCUMENTS

(Including Author, Title, Date, Pertinent Pages, Etc.)

|          |                 |  |
|----------|-----------------|--|
|          |                 |  |
|          |                 |  |
|          |                 |  |
| Examiner | Date Considered |  |

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP §609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to the applicant.

(19)



Europäisches Patentamt  
European Patent Office  
Office européen des brevets



(11) Publication number:

**0 631 250 A2**

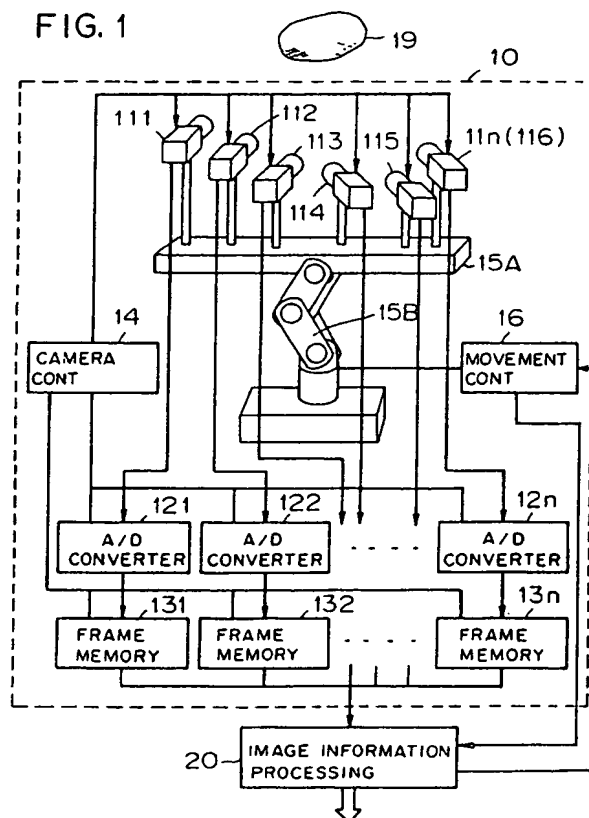
(12)

**EUROPEAN PATENT APPLICATION**(21) Application number: **94109510.1**(51) Int. Cl.<sup>5</sup>: **G06F 15/72**(22) Date of filing: **20.06.94**(30) Priority: **21.06.93 JP 149334/93**(43) Date of publication of application:  
**28.12.94 Bulletin 94/52**(84) Designated Contracting States:  
**DE FR GB**(71) Applicant: **NIPPON TELEGRAPH AND  
TELEPHONE CORPORATION**  
**1-6 Uchisaiwaicho 1-chome**  
**Chiyoda-ku**  
**Tokyo (JP)**(72) Inventor: **Mori, Takeaki**  
**Gurin Haitzu 7-5-301,**  
**510, Tsukui**  
**Yokosuka-shi,****Kanagawa (JP)****Inventor: Suzuki, Satoshi**  
**1-50-12, Tomioka-Nishi,**  
**Kanazawa-ku****Yokohama-shi,**  
**Kanagawa (JP)****Inventor: Yasuno, Takayuki**  
**968-12-2-101, Mutsuura,**  
**Kanazawa-ku**  
**Yokohama-shi,**  
**Kanagawa (JP)**(74) Representative: **Hoffmann, Eckart**  
**Patentanwalt,**  
**Blumbach & Partner,**  
**Bahnhofstrasse 103**  
**D-82166 Gräfelfing (DE)**(54) **Method and apparatus for reconstructing three-dimensional objects.**

(57) One of images of an object (19) captured by at least three cameras (111-11n) is defined as a basic image, and a basic backprojection line is defined which passes through a feature point in the basic image corresponding to a noted three-dimensional feature point on the object and the optical center of the camera having captured the basic image. Reference backprojection lines are calculated respective feature points on an epipolar line, obtained by projecting the basic backprojection line onto the image by each of the other cameras, and the optical center of said each camera. The coordinates of intersection points of the basic backprojection line and the reference backprojection lines are calculated and the number of the reference backprojection lines intersecting at each intersection point on the basic backprojection line is counted. The point having the maximum number of intersections is determined to be the position of the noted three-dimensional feature point. Similar processing is performed for all of the feature points in the basic image, by which the positions of corresponding three-dimensional feature points of the object are determined.

**EP 0 631 250 A2**

FIG. 1



## BACKGROUND OF THE INVENTION

The present invention relates to measurements of the position, shape and movement of a three-dimensional moving object and, more particularly, to a three-dimensional information reconstruction or recovery method and apparatus which can be used in the fields of three-dimensional information reconstruction, recognition and description (CG) of moving objects.

Conventional three-dimensional information reconstruction techniques can be classified into three categories. A first technique is stereopsis. This technique is one that establishes correspondence of points or lines between two right and left images taken by two cameras and estimates, from the positions of the cameras and pairs of corresponding points or lines on the right and left images, the positions of points and lines on a scene space corresponding to the pairs of corresponding points or lines. A second technique is a three-dimensional information recovery method using a moving camera. This is a method which tracks individual feature points on a number of images picked up by a moving camera and estimates the positions of points on a scene space corresponding to the feature points. A third technique is a backprojection method, which recovers or reconstructs a three-dimensional structure of an object in a scene space by projecting back feature points in images to the scene space.

With the first technique (see, for example, Kanade T., Okutomi T. and Nakahara M., "A multiple baseline stereo method," Proc. Image understanding Workshop, pp. 409-426, 1992 or United States Patent No. 4,654,872), many points on the surface of the object tend to be occluded from the sight line of either one of the cameras because of uneven object surface, and hence accurate positions of the corresponding feature points between the right and left images cannot be obtained, making it hard to obtain highly accurate three-dimensional information. The second technique (see, for example, Bolles R. C., Baker H. H. and Marimont D. H.: "Epipolar-plane image analysis: an approach to determining structure from motion," IJCV, Vol. 1, No. 1, pp. 7-55, 1987) cannot be applied to a moving object, because the object needs to stand still during its image capturing session by a moving camera. Recently, there has been proposed a technique which permits simultaneous extraction of the three-dimensional shape and motion of an object from many images taken by a single camera (see Tomasi C. and Tanade T.: "Shape and motion from image streams under orthography: a factorization method," IJCV, Vol. 9, No. 2, pp. 137-154, 1992). This technique recovers three-dimensional information basically by tracking feature points between the images; therefore, this method cannot obtain accurate three-dimensional information because the surface of the object partly occluded from the sight line of the camera while the camera or the object is moving. The reason for this is that since a noted feature point on the images repeatedly gets out of and into the view field, the locus of the feature points on the images frequently breaks, introducing difficulty in tracking the feature point. Hence, this technique is not suitable for use with a moving object. A third technique is a silhouette projection method (see, for example, Ahuja N. and Veenstra J.: "Generation octree from object silhouettes in orthographic views," IEEE Trans. PAMI, Vol. 11, No. 2, pp. 137-149, 1989). With this method, however, it is very difficult to acquire accurate three-dimensional information, because the generation of silhouette images is extremely difficult and unstable. Another example of the third technique is a method which recover or reconstruct edges of a three-dimensional object by extracting edges on images and giving a voting to the scene space through use of the extracted edges (see, for example, Hamano T., Yasuno T. and Ishii K.: "Direct estimation of structure from non-linear motion by voting algorithm without tracking and matching," Proc. of ICPR, Vol. 1, pp. 505-508, 1982 and S. Kawato: "3D Shape Recovering by Octree Voting Technique," PROCEEDINGS of SPLE-The International Society for Optical Engineering, 15-16 Nov. 1992). With such a method, however, since a plurality of feature points are simultaneously extracted, processes for the respective feature points interfere with each other, incurring possibility of a false feature point being extracted. A large number of images are needed to solve this problem. For a moving object, however, much time is consumed to take many images by one camera and a simultaneous image pickup system using many cameras is very expensive.

## SUMMARY OF THE INVENTION

It is therefore an object of the present invention to provide a method and apparatus which permit the acquisition or recovery of three-dimensional information of an object from its images taken by a small number of cameras.

A three-dimensional information recovery method according to a first aspect of the invention comprises the steps:

(a) wherein an object is imaged by  $n$  cameras to obtain  $n$  frames of images and these images are output as image data,  $n$  being an integer equal to or greater than 3;

(b) wherein the image data is received and feature points of the object on the  $n$  frames of images are respectively extracted;

(c) wherein one of the  $n$  cameras is defined as a basic camera, the image taken by the basic camera as a basic image, the other remaining cameras as reference cameras, the images taken by the reference cameras as reference images, and a backprojection line, which passes through the optical center of the basic camera and selected one of feature points in the basic image chosen in correspondence to a noted three-dimensional feature point on the object, is defined as a basic backprojection line;

(d) wherein the basic backprojection line is projected onto the reference images, thereby defining epipolar lines on the reference images, respectively;

(e) wherein backprojection lines, which pass through feature points on the epipolar lines on the reference images and the optical centers of the reference cameras, are defined as reference backprojection lines;

(f) wherein the coordinates of intersections of the basic backprojection line and the reference backprojection lines are respectively calculated, the numbers of intersections of the basic backprojection line and the reference backprojection lines at the respective intersections are counted, and the intersection of the maximum count value is determined to be the position of the above-mentioned noted three-dimensional feature point; and

(g) wherein a sequence of steps (c) through (f) are repeated for each feature point on the basic image to obtain the positions of the respective three-dimensional feature points on the object as its three-dimensional information.

A method according to a second aspect of the invention comprises the steps:

(a) wherein an object is imaged by  $n$  cameras to obtain  $n$  frames of images and these images are output as image data,  $n$  being an integer equal to or greater than 3;

(b) wherein the image data is received and feature points on the object on the  $n$  frames of images are respectively extracted;

(c) wherein one of the  $n$  cameras is defined as a basic camera, the image taken by the basic camera as a basic image, the other remaining cameras as reference cameras, the images taken by the reference cameras as reference images, and a backprojection line, which passes through the optical center of the basic camera and a feature point in the basic image chosen in correspondence to a noted three-dimensional feature point on the object, is defined as a basic backprojection line;

(d) wherein the basic backprojection line is projected onto the reference images, thereby defining epipolar lines on the reference images, respectively;

(e) wherein backprojection lines, which pass through feature points on the epipolar lines on the reference images and the optical centers of the reference cameras, are defined as reference backprojection lines;

(f) wherein the coordinates of intersections of the basic backprojection line and the reference backprojection lines are respectively calculated, the distribution of the intersections along the basic backprojection line is filtered by convolution to emphasize the concentration of the distribution of the intersections, and the position where the filtered distribution of the intersection is maximum is determined to be the position of the above-mentioned noted three-dimensional feature point; and

(g) wherein a sequence of steps (c) through (f) are repeated for each feature point on the basic image to obtain the positions of the respective three-dimensional feature points on the object as its three-dimensional information.

A three-dimensional information recovery apparatus according to the present invention comprises:

$n$  camera means disposed at different positions, for capturing images of an object and for outputting the images as image data,  $n$  being an integer equal to or greater than 3; and

image information processing means supplied with the image data from the  $n$  camera means, for acquiring or recovering three-dimensional information of the object from feature points in the respective images;

the image information processing means comprising:

feature extracting means for extracting feature points in the images;

basic backprojection line generating means wherein one of the  $n$  cameras is defined as a basic camera, the image taken by the reference camera as a basic image, the other remaining cameras as reference cameras and the images taken by the reference cameras as reference images and whereby a backprojection line, which passes through the optical center of the basic camera and a feature point in the basic image chosen in correspondence to the above-said noted three-dimensional feature point on the object, is obtained as a basic backprojection line;

epipolar line generating means for projecting the basic backprojection line onto the reference images to obtain epipolar lines;

reference backprojection line generating means for obtaining, as reference backprojection lines,

backprojection lines which pass through feature points on the epipolar lines on the reference images and the optical centers of the reference cameras;

three-dimensional feature extracting means for calculating the coordinates of intersections of the basic backprojection line and the reference backprojection lines, for counting the numbers of intersections of the basic backprojection line and the reference backprojection lines at the respective intersections and for determining the intersection of the maximum count value to be the position of the noted three-dimensional feature point;

whereby the positions of respective three-dimensional feature points on the object, corresponding to respective feature points on the basic image, are obtained as three-dimensional information of the object.

10 An apparatus according to a fourth aspect of the present invention comprising:

n cameras means disposed at different positions, for capturing images of an object and for outputting the images as image data, n being an integer equal to or greater than 3; and

image information processing means supplied with the image data from the n camera means, for acquiring or recovering three-dimensional information of the object from feature points in the respective images;

the image information processing means comprising:

feature extracting means for extracting feature points in the images;

basic backprojection line generating means wherein one of the n cameras is defined as a basic camera, the image taken by the basic camera as a basic image, the other remaining cameras as reference cameras, the images taken by the reference cameras as reference images and whereby a backprojection line, which passes through the optical center of the basic camera and a feature point in the basic image chosen in correspondence to the above-said noted three-dimensional feature point on the object, is obtained as a basic backprojection line;

epipolar line generating means for projecting the basic backprojection line onto the reference images to obtain epipolar lines;

reference backprojection line generating means for obtaining, as reference backprojection lines, backprojection lines which pass through feature points on the epipolar lines on the reference images and the optical centers of the reference cameras;

30 filtering means which calculates the coordinates of intersections of the basic backprojection line and the reference backprojection lines, respectively, and filters the distribution of the intersections along the basic backprojection line by convolution to emphasize the concentration of the distribution of the intersections; and

three-dimensional feature extracting means whereby the position where the filtered distribution of intersection is maximum is determined to be the position of the noted three-dimensional feature point;

35 whereby the positions of respective three-dimensional feature points on the object, corresponding to respective feature points on the basic image, are obtained as three-dimensional information of the object.

With the three-dimensional moving object information reconstructing methods and apparatuses according to the present invention, since the backprojection area is substantially limited by the introduction of a constraint of projective geometry, the interference of backprojection is minimized and three-dimensional information can be acquired by backprojection using a small number of images. Accordingly, the present invention permits reduction of the number of cameras needed to obtain ever-changing three-dimensional information of a moving object through simultaneous image capturing by a plurality of cameras.

#### BRIEF DESCRIPTION OF THE DRAWINGS

45 Figure 1 is a block diagram illustrating the configuration of an embodiment of the three-dimensional information reconstructing apparatus according to the present invention;

Figure 2 is a processing block diagram for explaining in detail an image information processing part 20 in Fig. 1;

50 Figure 3 is a backprojection diagram for explaining the principles of the present invention;

Figure 4 is a three-dimensional graph for explaining the probability of reference backprojection lines intersecting at the same point;

Figure 5A is a perspective view for explaining a pixel quantization error of a feature point in an image;

Figure 5B is a perspective view for explaining an error of the optical center;

55 Figure 6A is a diagram showing the intersection distribution on a basic backprojection line;

Figure 6B is a graph showing an example of the characteristic of a filter; and

Figure 6C is a diagram showing an example of a filtered intersection distribution.



DESCRIPTION OF THE PREFERRED EMBODIMENTS

In Fig. 1 there is shown in block form an embodiment of the three-dimensional information reconstructing apparatus according to the present invention. The illustrated apparatus is made up of an image input part 10 and an image information processing part 20. The image input part 10 is composed of  $n$  ( $n$  being an integer equal to or greater than 3) image capturing parts (for example, TV cameras, hereinafter referred to simply as cameras) 111 through 11 $n$ , A/D converters 121 through 12 $n$ , frame memories 131 through 13 $n$ , a camera controller 14, a movable table 15A, a driver 15B and a movement controller 16. The cameras 111 through 11 $n$  are mounted on the common moving table 15A at different heights. The moving table 15A is designed to be movable three-dimensionally by the driver 15A formed by a robot arm under the control of the movement controller 16.

The cameras 111 through 11 $n$  simultaneously capture images of an object 19 at a fixed period (or frame period) and each output an analog image signal. When the object 19 is not moving, its images need not always be captured simultaneously. The analog signals from the cameras 111 through 11 $n$  are converted by the A/D converters 121 through 12 $n$  into digital signals, which are stored, as image data for each frame, in the frame memories 131 through 13 $n$ . The operations of these cameras, A/D converters and frame memories are placed under the control of the camera controller 14. The image information processing part 20 is supplied with the image data read out of the frame memories 131 through 13 $n$  and obtains three-dimensional information of the object 19 by performing a process described below with reference to Fig. 2.

Figure 2 shows, in the form of functional blocks, the procedure of processing by the image information processing part 20 formed by a computer. Position and attitude information of the cameras at the time of inputting images of the object 19 by the cameras 111 through 11 $n$  is inputted into a camera information input part 21 of the image information processing part 20 from the movement controller 16. On the other hand, image data read out of the frame memories 131 through 13 $n$  are fed to a feature extracting part 22. The feature extracting part 22 extracts feature points (such as isosbestic points and regions of normal images and isosbestic points and regions of thermal images) from the  $n$  frames of images inputted from the image input part 10. A basic backprojection line generating part 23 selects, as a basic image, an arbitrary one of the  $n$  frames of images from which the feature points have been extracted and generates a basic backprojection line which passes through a feature point on the basic image and the optical center of the camera having taken the basic image inputted from the camera information input part 21 (which camera will hereinafter be referred to as a basic camera). An epipolar line generating part 24 provides epipolar lines which are generated by projecting the above-mentioned basic backprojection line onto images (reference images) taken by cameras (reference cameras) other than the basic camera. A reference backprojection line generating part 25 provides all reference backprojection lines which pass through feature points on the epipolar lines on the reference images or within a certain distance range on either side of the epipolar lines and optical centers of the reference cameras corresponding to the epipolar lines. An intersection extracting part 26 calculates the coordinates of all intersections of the basic backprojection line generated by the basic backprojection line generating part 23 and the reference backprojection lines generated by the reference backprojection line generating part 25. An intersection counting part 27 counts the number of intersections at intersecting points or regions of the basic backprojection line and the reference backprojection lines extracted by the intersection extracting part 26 along the basic backprojection line. A filtering part 28 filters, along the basic backprojection line, the intersection count distribution provided by the intersection counting part 27. A three-dimensional feature extracting part 29 determines, as the position of a three-dimensional feature point, one of the intersections which is closest to the peak position where the filtered intersection distribution exceeds a fixed threshold value. A three-dimensional movement extracting part 2A calculates movement of each three-dimensional feature point and region on the basis of each three-dimensional feature point and region information at every time interval extracted by the three-dimensional feature extracting part 29.

Next, the operation of this invention apparatus of such a construction will be described with reference to Figs. 1, 2 and 3. In Fig. 3 inverted images formed on focal planes of the cameras 111 through 11 $n$  disposed as depicted are imaginarily projected back forwardly of the cameras and shown, for convenience's sake, as erect images 32. The camera controller 14 simultaneously activates the  $n$  cameras 111 through 11 $n$  from an input time  $t$ , converts the resulting analog image signals by the A/D converters 121 through 12 $n$  into digital image signals and captures them into the  $n$  frame memories 131 through 13 $n$  to obtain image data of one frame in each memory. The camera information input part 21 obtains camera information at the time  $t$  through calibration (see, for example, J. Weng and P. Cohen and M. Herniou: "Calibration of stereo cameras using a non-linear distortion model," Proc. ICCV, pp. 246-253, 1990).

By the feature extracting part, feature points on each image 32 (for example, isosbestic points and regions in an image, contour lines and regions in a differential image and isothermal points and regions in a thermal image) are extracted. In this example, edges of the image are extracted feature points by a Canny filter (Canny J. F.: "A computational approach to edge detection," IEEE Trans. PAMI, Vol. 8, No. 6, pp. 679-698, 1986). Next, the basic backprojection generating part 23 selects an image 32 of an arbitrary one of the n frames as a basic image 32s and calculates a backprojection line 34s which passes through one noted feature point on the basic image 32s and the optical center 31s of the camera (that is, the basic camera) corresponding to the image 32s which is inputted from the camera information input part 21. The coordinates  $(x_b, y_b, z_b)$  of an arbitrary point on the backprojection line 34s satisfy the following equation:

$$\frac{x_b - x_c}{x_p - x_c} = \frac{y_b - y_c}{y_p - y_c} = \frac{z_b - z_c}{z_p - z_c} \quad (1)$$

where  $(x_p, y_p, z_p)$  and  $(x_c, y_c, z_c)$  are coordinates of the noted feature point on the image 32s and the optical center 31s of the basic camera, respectively. Such a backprojection line 34s will hereinafter be referred to as a basic backprojection line. An arbitrary point  $E = (x_e, y_e)$  on an epipolar line, which is obtained by projecting the basic backprojection line onto each reference image, has such relations as shown below.

$$E = R^T (B' - T') \quad (2)$$

$$R^T = \begin{pmatrix} r_{11} & r_{21} \\ r_{12} & r_{22} \end{pmatrix}$$

$$T' = (t_1, t_2)^T$$

where  $R^T$  and  $T'$  are a two-dimensional expression of transposed version of a matrix which defines the rotation of the camera optical axis with respect to the coordinate system and a two-dimensional expression of a vector which defines the translation of the camera optical axis with respect to the coordinate system, respectively, and  $B' = (x_b, y_b)^T$  is a two-dimensional expression of an arbitrary point on the backprojection line. As is evident from Fig. 3, the basic backprojection line 34s is present in a plane containing the optical center 31 of every reference camera and the epipolar line 35 by the reference camera. This plane will hereinafter be called an epipolar plane by the reference camera. Three-dimensional feature points on the object 19 present in an extension plane of the epipolar plane of each camera are all projected onto the epipolar line 35 of the camera. Next, the reference backprojection line generating part 25 detects all feature points 33 on the above-mentioned epipolar line 35. One of these feature points 33e on the epipolar line 35 corresponds to the noted feature point 33s. In practice, however, there are cases where these feature points are not on the epipolar line because of an image quantization error. Accordingly, feature point within a fixed distance from the epipolar line 35 are also regarded as its feature points 33e. This is followed by obtaining backprojection lines 34 which pass through the feature points 33e on the epipolar line 35 in the image 32 by each reference camera and the optical center 31 of the reference camera. These backprojection lines 34 will hereinafter be referred to as reference backprojection lines.

Next, the intersection extracting part 26 calculates the coordinates of intersections of the basic backprojection line 34s and the reference backprojection lines 34. Owing to the presence of a feature extraction error on each image 32 and an error in the optical center of each camera, the situation occasionally arises where the basic backprojection line 34s and the reference backprojection lines 34 do not intersect in the three-dimensional space. In this embodiment, the shortest distance between the basic backprojection line 34s and each reference backprojection line 34 is calculated and if this shortest distance is smaller than a fixed threshold value, the basic backprojection line 34s and the reference backprojection line 34 concerned are regarded as intersecting each other. That is, if the rank of the following matrix with two arbitrary points  $B_{b1} = (x_{b1}, y_{b1}, z_{b1})^T$  and  $B_{b2} = (x_{b2}, y_{b2}, z_{b2})^T$  on the basic backprojection line 34s and two arbitrary points  $B_{r1} = (x_{r1}, y_{r1}, z_{r1})^T$  and  $B_{r2} = (x_{r2}, y_{r2}, z_{r2})^T$  on the reference backprojection line 34 is 1, these two lines do not extend in parallel to each other.

$$\begin{pmatrix} x_{b2} - x_{b1} & y_{b2} - y_{b1} & z_{b2} - z_{b1} \\ x_{r2} - x_{r1} & y_{r2} - y_{r1} & z_{r2} - z_{r1} \end{pmatrix} \quad (3)$$

In this instance, two points  $(x_{bs}, y_{bs}, z_{bs})^T$  and  $(x_{rs}, y_{rs}, z_{rs})^T$ , which stay on the basic backprojection line 34s and the reference backprojection line 34, respectively, and provide the shortest distance between the two lines, are given as follows:

$$\begin{bmatrix} x_{bs} \\ y_{bs} \\ z_{bs} \end{bmatrix} = \begin{bmatrix} x_{b1} \\ y_{b1} \\ z_{b1} \end{bmatrix} + K_b \begin{bmatrix} x_{b2} - x_{b1} \\ y_{b2} - y_{b1} \\ z_{b2} - z_{b1} \end{bmatrix} \quad (4)$$

$$\begin{bmatrix} x_{rs} \\ y_{rs} \\ z_{rs} \end{bmatrix} = \begin{bmatrix} x_{r1} \\ y_{r1} \\ z_{r1} \end{bmatrix} + K_r \begin{bmatrix} x_{r2} - x_{r1} \\ y_{r2} - y_{r1} \\ z_{r2} - z_{r1} \end{bmatrix} \quad (5)$$

where:

$$K_b = (G_r H - F_r G_b) / (F_b F_r - H^2)$$

$$K_r = (F_b G_r - G_b H) / (F_b F_r - H^2)$$

$$F_b = (x_{b2} - x_{b1})^2 + (y_{b2} - y_{b1})^2 + (z_{b2} - z_{b1})^2$$

$$F_r = (x_{r2} - x_{r1})^2 + (y_{r2} - y_{r1})^2 + (z_{r2} - z_{r1})^2$$

$$G_b = (x_{b2} - x_{b1})(x_{b1} - x_{r1}) + (y_{b2} - y_{b1})(y_{b1} - y_{r1}) + (z_{b2} - z_{b1})(z_{b1} - z_{r1})$$

$$G_r = (x_{r2} - x_{r1})(x_{r1} - x_{b1}) + (y_{r2} - y_{r1})(y_{r1} - y_{b1}) + (z_{r2} - z_{r1})(z_{r1} - z_{b1})$$

$$H = (x_{b2} - x_{b1})(x_{r1} - x_{r2}) + (y_{b2} - y_{b1})(y_{r1} - y_{r2}) + (z_{b2} - z_{b1})(z_{r1} - z_{r2})$$

If the distance  $d$  between these two points satisfies the condition that it be smaller than a threshold value  $\lambda$  as shown below,

$$d = \{(x_{bs} - x_{rs})^2 + (y_{bs} - y_{rs})^2 + (z_{bs} - z_{rs})^2\}^{1/2} < \lambda \quad (6)$$

the basic backprojection line 34s and the reference backprojection line 34 on which these two point stay, respectively, are regarded as intersecting with each other, and the coordinates of the position of the intersection 36 are determined to be the midpoint of the distance between the two points.

$$x = (x_{bs} + x_{rs})/2$$

$$y = (y_{bs} + y_{rs})/2$$

$$z = (z_{bs} + z_{rs})/2 \quad (7)$$

Next, the intersection counting part counts the number of intersections of the basic backprojection line 34s and all of the reference backprojection lines 34.

Incidentally, in the case of using images taken by two cameras ( $n = 2$ ), there is the possibility of a plurality of reference backprojection lines 34 intersecting one basic backprojection line 34s at a plurality of positions. In such an instance, it cannot uniquely be determined to which intersection the three-dimensional feature point corresponds, the latter corresponding to the feature point 33s through which the basic backprojection line 34s passes. Now, description will be given of this problem.

In the interests of clarity, let the number of feature points on the epipolar line of each reference camera obtained by projecting the basic backprojection line 34s onto each reference image be represented by  $M$ . Only one or none of the  $M$  feature points on the epipolar line on the image by each camera correspond to the noted feature point. Now, assume that one of the  $M$  feature point on each epipolar line corresponds to the noted feature point. Accordingly, one reference camera provides  $M$  reference backprojection lines and the probability that any one of the  $M$  intersections of  $M$  reference backprojection lines and the basic backprojection line corresponds to a noted feature point is  $\beta = 1/M$ . The same goes for the other reference cameras. Hence, the probability  $\alpha$  that one of the reference backprojection lines from every reference camera intersects the basic backprojection line 34s at any intersection thereon is  $\alpha = \beta^{n-1}$ . When only one feature point is present on the epipolar line of the reference camera (that is, when  $M = 1$ ), that feature point

corresponds to the noted feature point and the reference backprojection line intersects the noted feature point; in this case, no problem arises. It is when the number  $M$  of reference backprojection lines from each reference camera is 2 or more that matters. Therefore, discussion will be made below on the case where  $\beta = 1/M$  is in the range of between 0 to 0.5.

Figure 4 is a graph showing the probability  $\alpha$  that  $(n - 1)$  reference backprojection lines from  $(n - 1)$  reference cameras intersect at one point on the basic backprojection line when the value  $\beta$  is in the range of  $0 < \beta \leq 0.5$ . As is evident from the graph, even in the case where  $\beta = 0.5$ , if the number  $n$  of all cameras (i.e. the number of viewing points) is equal to or larger than 6, the probability  $\alpha$  of five reference backprojection lines from reference cameras intersecting at one point on the basic backprojection line is as negligibly small as  $(0.5)^5 = 0.03$ . Also when  $n = 5$  (the number of reference cameras is four), the value  $\alpha$  is as small as 0.06. That is, the possibility that a plurality of reference backprojection lines from respective reference cameras happen to intersect at the noted feature point decreases rapidly as the number  $n$  of cameras increases. Conversely speaking, when the number of reference backprojection lines which actually intersect at a certain point on the basic backprojection line is plural (four or more, for instance), it implies that the intersection is very likely to be a noted feature point. Then, it is possible to uniquely determine or locate the position of the noted three-dimensional feature point by counting in the intersection counting part 27 the number of reference backprojection lines intersecting at each point on the basic backprojection line 34s (hereinafter referred to simply as the intersection number) and then detecting in the three-dimensional feature point extracting part 29 the position providing the maximum intersection number, as described previously. In this way, any given three-dimensional feature point on the object 19 can be made to correspond to a unique intersection on the basic backprojection line.

In view of the above, the value  $n$  is set to 3 or larger on the principles of the present invention, but in a preferred embodiment six different viewing points, that is, six cameras are used. Such a configuration will increase the possibility that a three-dimensional feature point on the object 19, occluded from a particular one of the viewing points or cameras, can also be observed by or seen from the other cameras and hence located.

In actual measurements, as shown in Figs. 5A and 5B, backprojection lines may sometimes deviate from their original directions owing to positional errors of the feature points 33 on the image 32 and of the optical center 31 of the camera. For example, since the image 32 composed of image data provided by each camera has been quantized by pixels of the image, the position of each feature point on the image contains a quantization error by the pixels. Now, suppose that the position  $P$  of the true feature point 33 on the image 32 by a reference camera deviates to  $P'$  due to the quantization error as depicted in Fig. 5A. Since the reference backprojection line 34 rotates and deviates about the optical center 31, it is determined that the true feature point  $B$  on the basic backprojection line 34s is at  $B'$  (not always on the basic backprojection line 34s). Alternatively, if the optical center determined by measurement deviates to  $L'$  from the true optical center  $L$  as depicted in Fig. 5B, then the true feature point  $B$  on the basic backprojection line 34s will be regarded as lying at  $B''$ .

Because of the above-mentioned deviation of the reference backprojection line 34 under the influence of such kinds of errors, the feature point 33 mapped onto each image from the same three-dimensional feature point on the object 19 and the reference backprojection line 34 passing through the optical center 31 corresponding to the feature point 33 are likely to intersect at a point other than the true position of the three-dimensional feature point. In such a situation, if the point  $B'$  or  $B''$  is within a threshold distance from the basic backprojection line 34s, it is likely to be counted at an intersection different from the position of the true noted feature point on the basic backprojection line 34s at the time of counting the intersection number by the intersection counting part 27. In this instance, if the position of the intersection which provides the maximum count value in excess of a predetermined value is directly extracted, as a three-dimensional feature point, by threshold value processing, there is a fear that a three-dimensional feature point other than the noted one is extracted as a pseudo or false three-dimensional feature point.

Since these errors of the backprojection lines scatter randomly in direction and magnitude, however, it can be estimated that the central region of the actually obtainable discrete and expansive count distribution, around which intersections concentrate, is most likely to include the true three-dimensional feature point. It is therefore possible to decrease the likelihood of extracting pseudo or false three-dimensional feature points by filtering the intersection count distribution along the basic backprojection line 34s by the filtering part 28 to emphasize the distribution at points where reference backprojection lines concentratedly intersect the basic backprojection line. To this end, it is possible to use a square filter, a triangular filter, a Gaussian filter, a Laplacian-Gaussian filter and so forth.

For the discrete distribution  $P(s)$  of intersection counts 51 along the basic backprojection line 34s, such as shown in Fig. 6A, the filtering process using a function  $f(t)$  can be done by a convolution based on the

following equation.

$$P'(s) = \int P(s-t) \cdot f(t) dt \quad (7)$$

- 5 where  $P(s)$  and  $P'(s)$  are the distributions of backprojection line intersection counts before and after the convolution of the function  $f(t)$ , respectively, and  $s$  is the position along the basic backprojection line 34s. The filter function  $f(t)$  can be obtained with such various filters as mentioned above. In this embodiment, a Laplacian-Gaussian filter (commonly called a  $\nabla^2 G$  filter) 52, which has such a characteristic as shown in Fig. 6B, is used to perform the convolution with the distribution of the intersection counts 51 along the basic  
10 backprojection line 34s. Since the  $\nabla^2 G$  filter has a property of emphasizing peaks of a continuous waveform while suppressing high-frequency discrete noise, the convolution suppresses discrete intersection counts and emphasizes peaks of highly concentrated intersection counts as shown in Fig. 6C. Thus, the extraction of false three-dimensional feature point can be prevented by detecting peaks 51p of the intersection counts on the basic backprojection line 34s in the three-dimensional feature extracting part 29 and then determin-  
15 ing that the intersection closest to the peak position concerned is the position of the noted three-dimensional feature point.

As described above, intersections of the basic backprojection line 34s and the reference backprojection lines 34 from the reference cameras are defined on the basic backprojection line 34s and the number of intersections at each intersection is counted. In practice, however, owing to the afore-mentioned optical  
20 center error or pixel quantization error at the feature point, the intersection of one of the reference backprojection lines 34 from every reference camera and the basic backprojection line 34s, corresponding to the true noted three-dimensional feature point, does not become one in the strict sense. Then, in the case of counting the number of intersections at each point of intersection, it is necessary to divide the basic backprojection line 34s lengthwise thereof every predetermined small unit length  $\Delta s$  and count the  
25 intersection number in each divided unit length  $\Delta s$ . In this case, since the resolution for the intersections increases with a decrease in the unit length  $\Delta s$ , the number of intersections present in the same divided unit of length  $\Delta s$  on the basic backprojection line 34s will decrease. When the unit length  $\Delta s$  is selected small to such an extent that all intersections are separated, the maximum intersection number at each intersection goes one and the noted three-dimensional feature point cannot uniquely be identified. Also in  
30 this instance, however, intersections cluster about the noted three-dimensional feature point as shown in Fig. 6A, for example; hence, the filtering convolution for the discrete distribution of intersections, such as described above, provides a peak at the center of the intersection clustering region, making it possible to determine that the intersection closest to the peak position is the position of the true three-dimensional feature point.

- 35 By repeating the above-described process for each of the feature points 33s on the basic image 32s, the coordinates of all three-dimensional feature points 36t corresponding to all the feature points 33s on the basic image 32s can be determined. In this case, the backprojection lines which intersect at the position determined to correspond to the noted three-dimensional feature point are related only to the noted three-dimensional feature point. Consequently, by removing the feature points 33 and 33s on the images 32 and  
40 32s corresponding to the three-dimensional feature points determined immediately before obtaining other three-dimensional feature points, it is possible to reduce the interference of backprojection and the computational cost. Furthermore, by selecting another image as a basic one and repeating the above-described process until all images are selected, it is possible to obtain the positions of three-dimensional feature points corresponding to feature points on images obtained from all of the viewing points, that is,  
45 taken by all of the cameras used.

Finally, the three-dimensional movement extracting part 2A extracts three-dimensional movement information of the moving object by calculating temporal changes of the three-dimensional feature point information available from the three-dimensional feature extracting part 29. A concrete method therefor can be implemented by a three-dimensional extended version of a two-dimensional movement information  
50 extracting method (see, for example, R. Agarwal and J. Sklansky: "Optical flow estimation by the clustering of constraint sets in velocity space," Technical Report TP-91-2, Department of Electrical and Computer Engineering, Pattern Recognition and Image Modelling Project, University of California, Irvine, CA, USA).

Moreover, it is also possible to carry out tracking of a moving object by repeating operations of: inputting the three-dimensional movement information from the three-dimensional movement extracting part  
55 2A into the movement controller 16 of the image input part 10; planning the next viewing point; activating the driver 15 to move the plurality of image capturing parts 111 through 11n and, at the same time, measuring the movement of the driver 15 at an input time  $t + w$ ; and inputting the measured value into the camera information input part.

As described above, according to the present invention, since the backprojection area is substantially limited by the introduction of a projective geometric constraint, the interference of backprojection is minimized, and consequently, three-dimensional information can be obtained by backprojection onto a small number of images. Hence, the ever-changing three-dimensional shape of a moving object can be recovered by simultaneous image pickup with a small number of cameras.

It will be apparent that many modifications and variations may be effected without departing from the scope of the novel concepts of the present invention.

## Claims

1. A method of obtaining three-dimensional information of an object by capturing images of said object by a plurality of cameras and making three-dimensional feature points on said object to correspond to feature points in said images, said method comprising the steps:
  - (a) wherein said object is imaged by  $n$  cameras to obtain  $n$  frames of images and said images are output as image data, said  $n$  being an integer equal to or greater than 3;
  - (b) wherein said image data is received and feature points of said object on said  $n$  frames of images are respectively extracted;
  - (c) wherein one of said  $n$  cameras is defined as a basic camera, the image picked up by said basic camera as a basic image the other remaining cameras as reference cameras, the images captured by said reference cameras as reference images, and a backprojection line, which passes through the optical center of said basic camera and a feature point in said basic image chosen in correspondence to a noted three-dimensional feature point on said object, is defined as a basic backprojection line;
  - (d) wherein said basic backprojection line is projected onto said reference images to define thereon epipolar lines, respectively;
  - (e) wherein backprojection lines, which pass through said feature points on said epipolar lines on said reference images and optical centers of said reference cameras, are defined as reference backprojection lines;
  - (f) wherein the coordinates of intersection points of said basic backprojection line and said reference backprojection lines are calculated, the numbers of intersections of said basic backprojection line and said reference backprojection lines at their respective intersection points are counted, and the intersection point of the maximum count value is determined to be the position of said noted three-dimensional feature point; and
  - (g) wherein said steps (c) through (f) are repeated for each of said feature points on said basic image to obtain the position of said three-dimensional feature points on said object as its three-dimensional information.
2. A method of obtaining three-dimensional information of an object by capturing images of said object by a plurality of cameras and making three-dimensional feature points on said object to correspond to feature point in said images, said method comprising the step:
  - (a) wherein said object is imaged by  $n$  cameras to obtain  $n$  frames of images and said images are output as image data, said  $n$  being an integer equal to or greater than 3;
  - (b) wherein said image data is received and feature points of said object on said  $n$  frame of images are respectively extracted;
  - (c) wherein one of said  $n$  cameras is defined as a basic camera, the image taken by said basic camera as a basic image, the other remaining cameras as reference cameras, the images captured by said reference cameras as reference images, and a backprojection line, which passes through the optical center of said basic camera and a feature point in said basic image chosen in correspondence to a noted three-dimensional feature point on said object, is defined as a basic backprojection line;
  - (d) wherein said basic backprojection line is projected onto said reference images to define thereon epipolar lines, respectively;
  - (e) wherein backprojection lines, which pass through said feature points on said epipolar lines on said reference images and optical centers of said reference cameras, are defined as reference backprojection lines;
  - (f) wherein the coordinates of intersection points of said basic backprojection line and said reference backprojection lines are calculated, the distribution of said intersection points along said basic backprojection line is filtered by convolution to emphasize the concentration of said distribution of

said intersection points, and the position where said filtered distribution of the intersection points is maximum is determined to be the position of said noted three-dimensional feature point; and  
 (g) wherein said steps (c) through (f) are repeated for each of said feature points on said basic image to obtain the position of said three-dimensional feature points on said object as its three-dimensional information.

3. The method of claim 1 or 2, wherein a position said basic backprojection line is defined every predetermined unit length of said basic backprojection line lengthwise thereof and the number of intersections at each of said intersection point is the number of intersections in said unit length at the position on said basic backprojection line through which said reference backprojection lines pass.

4. The method of claim 3, wherein said reference backprojection lines intersecting with said basic backprojection line are reference backprojection lines lying within a predetermined range of distance from said basic backprojection line.

5. The method of claim 1, wherein said step of counting said number of intersections is a step wherein the distribution of said intersections along said basic backprojection line is filtered by convolution to emphasize the concentration of said distribution of said intersections and the position where said filtered distribution of said intersections is maximum is determined to be the position of said noted three-dimensional feature point.

6. The method of claim 2 or 5, wherein said filtering is performed by a Laplacian-Gaussian filter.

7. The method of claim 2 or 5, wherein said filtering is performed by a square filter.

8. The method of claim 2 or 5, wherein said filtering is performed by a triangular filter.

9. The method of claim 2 or 5, wherein said filtering is performed by a Gaussian filter.

10. The method of claim 1 or 2, wherein said step (g) includes a step wherein each time said feature point corresponding to said noted three-dimensional feature point is identified by performing said series of steps (c) through (f), said identified feature point is removed from candidates for processing in the subsequent repetition of said steps.

11. The method of claim 1 or 2, wherein the number  $n$  of said cameras is in the range of from 5 to 8.

12. The method of claim 1 or 2, further comprising a step wherein upon each completion of said step (g), said step (g) is newly repeated for every one of said cameras, using one of said reference cameras as a new basic camera and said basic camera as a new reference camera.

13. The method of claim 1 or 2, further comprising a step wherein upon each repetition of simultaneous capturing of images of said object by said  $n$  cameras at a fixed period, three-dimensional information of said object is obtained following said series of steps.

14. A three-dimensional information reconstructing apparatus comprising:

$n$  camera means disposed at different positions, for capturing images of an object and for outputting said images as image data, said  $n$  being an integer equal to or greater than 3; and

image information processing means supplied with said image data from said  $n$  camera means, for acquiring three-dimensional information of said object from feature points in said images;

said image information processing means comprising:

feature extracting means for extracting said feature points in said images;

basic backprojection line generating means wherein one of said  $n$  camera means is defined as a basic camera, the image captured by said basic camera means as a basic image, the other remaining camera means as reference camera means, the images captured by said reference camera means as reference images, and whereby a backprojection line, which passes through the optical center of said basic camera means and a feature point in said basic image chosen in correspondence to a noted three-dimensional feature point on said object, is obtained as a basic backprojection line;

epipolar line generating means for projecting said basic backprojection line onto said reference

images to obtain epipolar lines;

reference backprojection line generating means for obtaining, as reference backprojection lines, backprojection lines which pass through feature points on said epipolar lines on said reference images and the optical centers of said reference camera means; and

5 three-dimensional feature extracting means for calculating the coordinates of intersection points of said basic backprojection line and said reference backprojection lines, for counting the number of intersections of said basic backprojection line and said reference backprojection lines at each of said intersection points and for determining that one of said intersection points which has the maximum count value is the position of said noted three-dimensional feature point;

10 whereby the positions of respective three-dimensional feature points on said object, corresponding to respective feature points on said basic image, are obtained as said three-dimensional information of said object.

15. A three-dimensional information reconstructing apparatus comprising:

15 n camera means disposed at difference positions, for capturing images of an object and for outputting said images as image data, said n being an integer equal to or greater than 3; and

image information processing means supplied with said image data from said n camera means, for acquiring three-dimensional information of said object from feature points in said images;

said image information processing means comprising:

20 feature extracting means for extracting feature points in said images;

basic backprojection line generating means wherein one of said n camera means is defined as a basic camera means, the image captured by said basic camera means as a basic image, the other remaining camera means as reference cameras, the images captured by said reference camera means as reference imaged, and whereby a backprojection line, which passes through the optical center of said basic camera means and a feature point in said basic image chosen in correspondence to a noted three-dimensional feature point on said object, is obtained as a basic backprojection line;

25 epipolar line generating means for projecting said basic backprojection line onto said reference images to obtain epipolar lines;

30 reference backprojection line generating means for obtaining, as reference backprojection lines, backprojection lines which pass through feature points on said epipolar lines on said reference images and the optical centers of said reference camera means;

35 filtering means which calculates the coordinates of intersection points of said basic backprojection line and said reference backprojection lines, respectively, and filters the distribution of said intersection points along said basic backprojection line by convolution to emphasize the concentration of said distribution of said intersection points; and

three-dimensional feature extracting means whereby the position where said filtered distribution of said intersection points is maximum is determined to the position of said noted three-dimensional feature point;

40 whereby the positions of respective three-dimensional feature points on said object, corresponding to respective feature points on said basic image, are obtained as said three-dimensional information of said object.

16. The apparatus of claim 14, wherein said feature extracting means includes filter means for emphasizing the concentration of said distribution of intersections by convolution along said basic backprojection line and the position where said filtered distribution of intersections is maximum is determined to be the position of said noted three-dimensional feature point.

17. The method of claim 15 or 16, wherein said filter means is a Laplacian-Gaussian filter.

50 18. The apparatus of claim 15 or 16, wherein said filter means is square filter.

19. The apparatus of claim 15 or 16, wherein said filter means is a triangular filter.

20. The apparatus of claim 15 or 16, wherein said filter means is a Gaussian filter.

55 21. The method of claim 14 or 15, wherein said number n of cameras means is in the range of from 5 to 8.



22. The apparatus of claim 14 or 15, further comprising camera control means for repeating simultaneous image capturing of said object by said n camera means at a fixed period and for supplying said image data to said image information processing means upon each generation of said image data.

5

10

15

20

25

30

35

40

45

50

55

FIG. 1

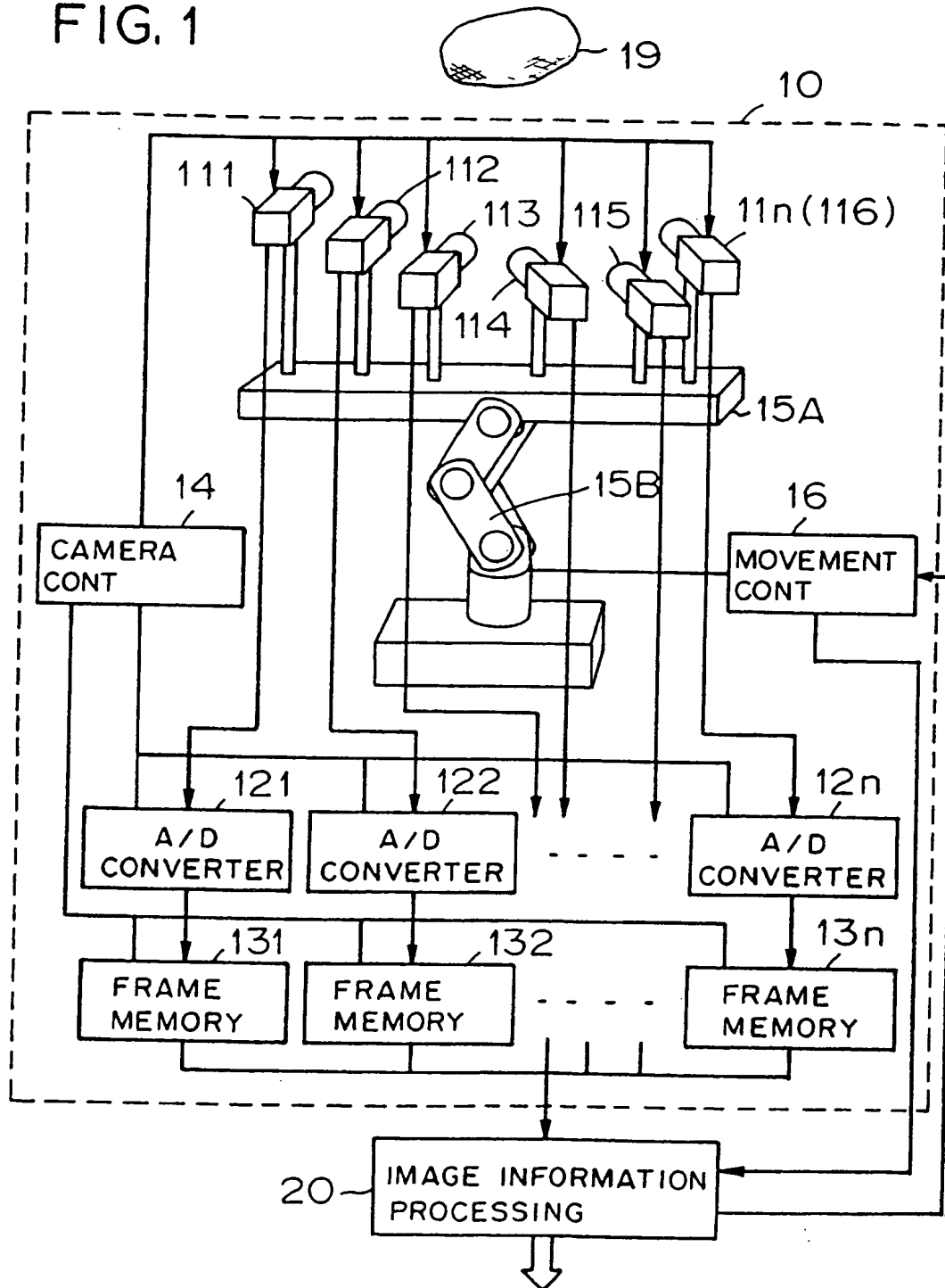


FIG. 2

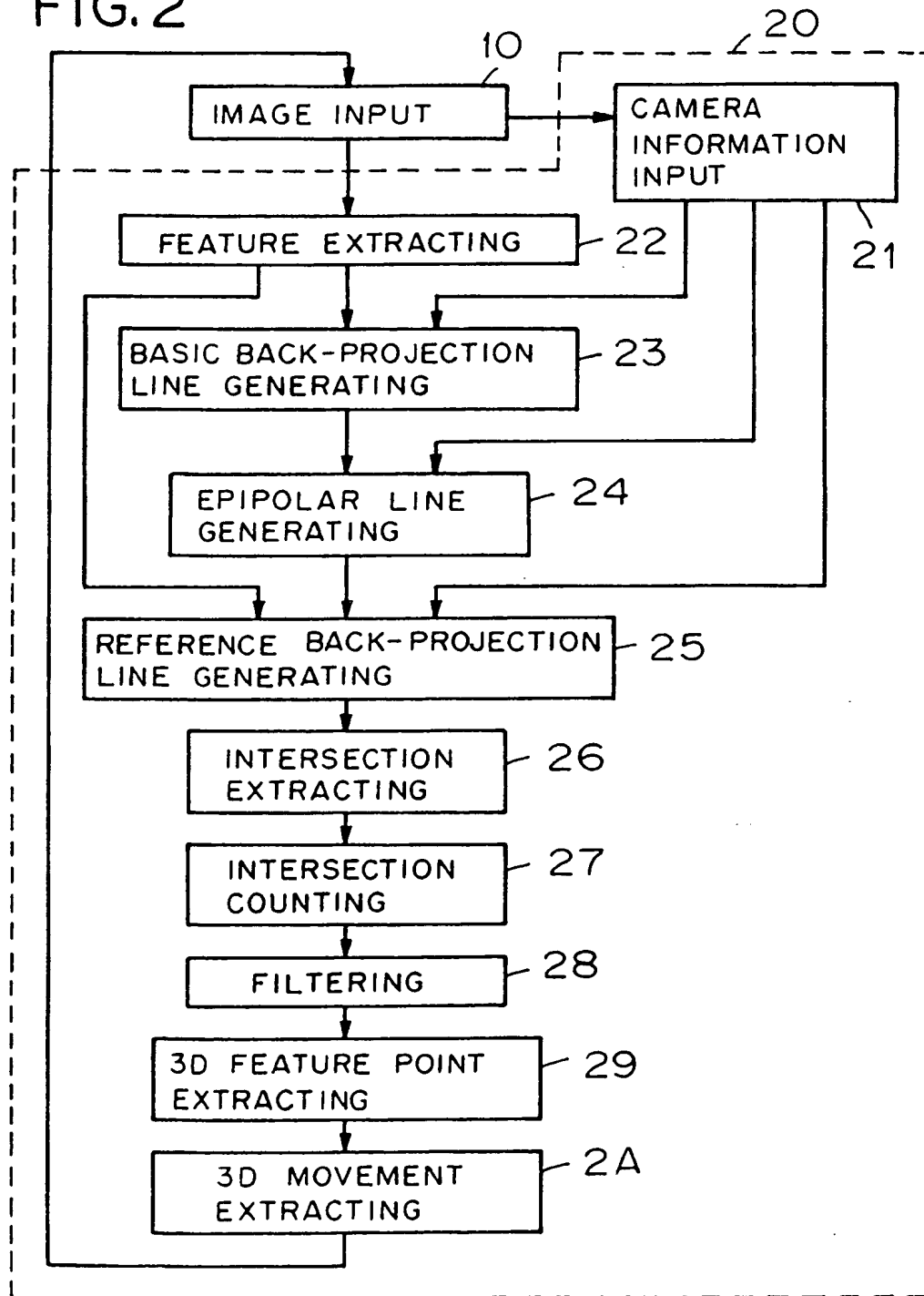


FIG. 3

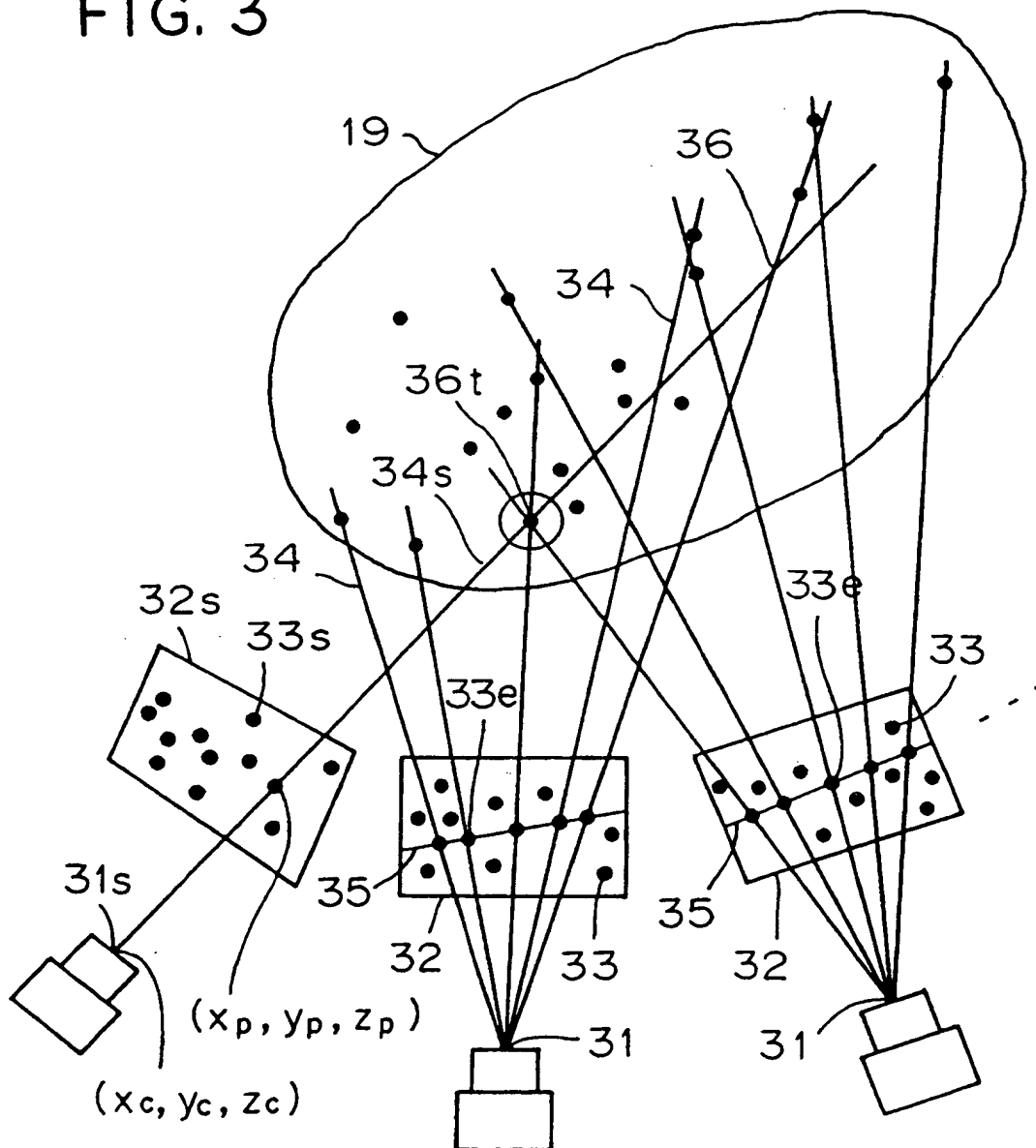


FIG. 4

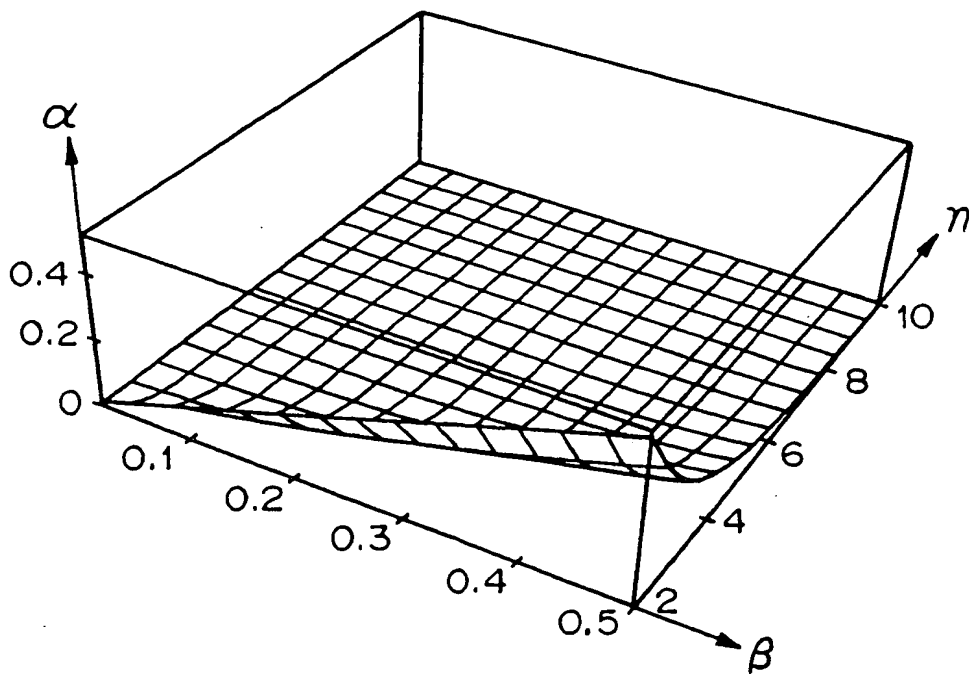


FIG. 5A

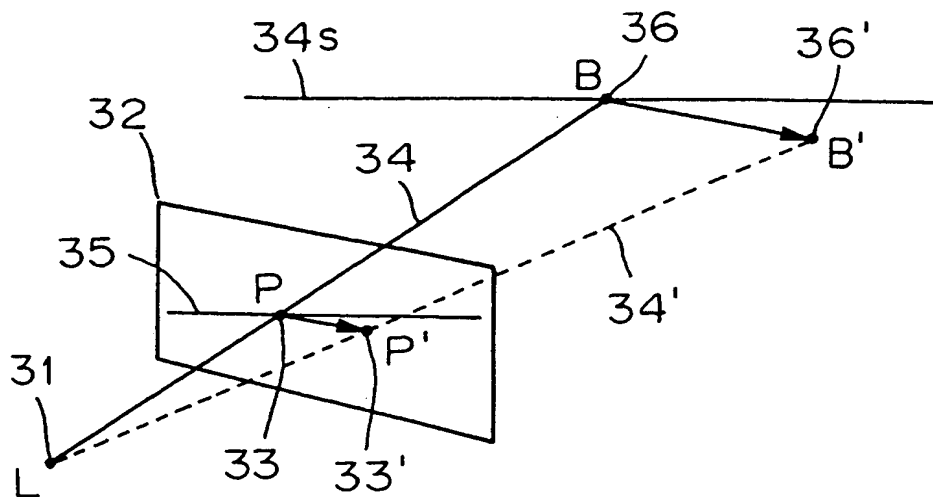


FIG. 5B

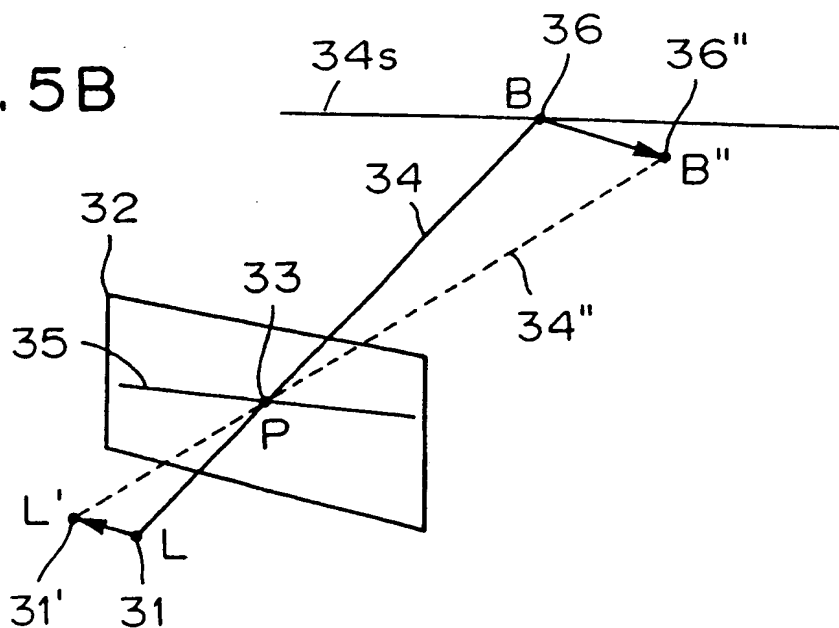


FIG. 6A

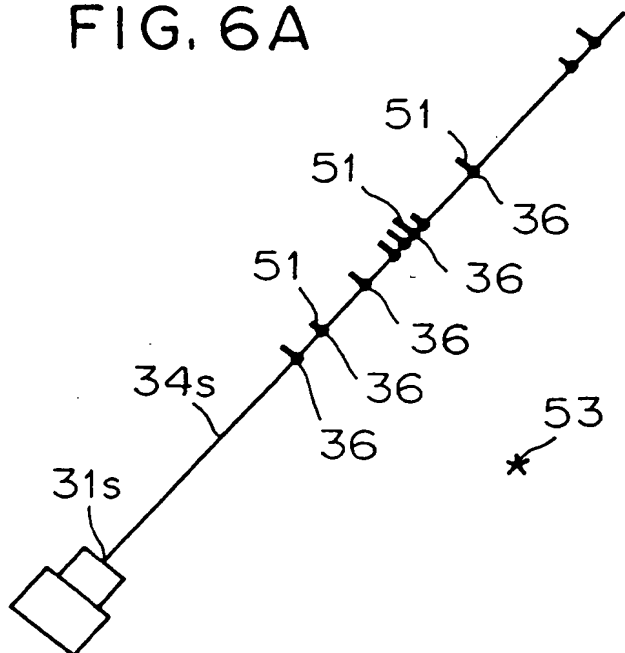


FIG. 6B

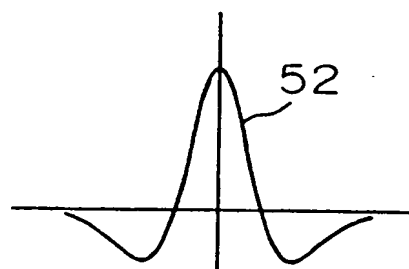
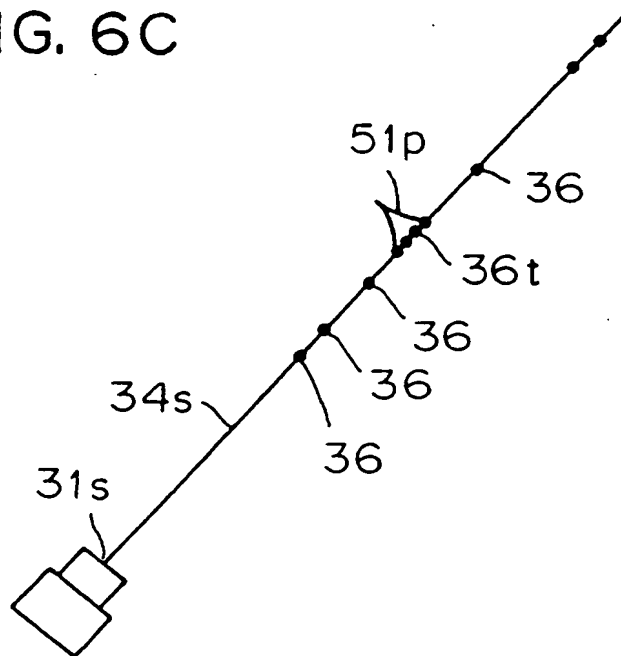
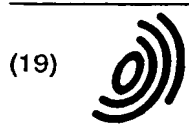


FIG. 6C





(19)

Europäisches Patentamt

European Patent Office

Office européen des brevets



(11)

EP 0 631 250 A3

(12)

## EUROPEAN PATENT APPLICATION

(88) Date of publication A3:  
03.01.1996 Bulletin 1996/01

(51) Int. Cl.<sup>6</sup>: G06F 15/72, G06F 15/70

(43) Date of publication A2:  
28.12.1994 Bulletin 1994/52

(21) Application number: 94109510.1

(22) Date of filing: 20.06.1994

(84) Designated Contracting States:  
DE FR GB

(30) Priority: 21.06.1993 JP 149334/93

(71) Applicant: NIPPON TELEGRAPH AND  
TELEPHONE CORPORATION  
Tokyo (JP)

(72) Inventors:

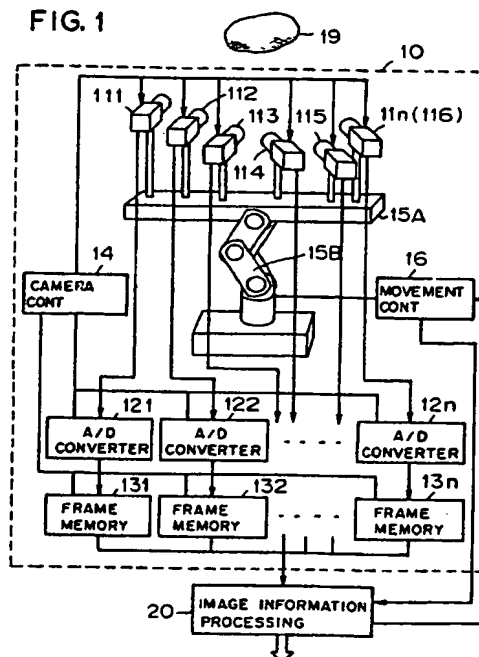
- Mori, Takeaki  
Yokosuka-shi, Kanagawa (JP)
- Suzuki, Satoshi  
Yokohama-shi, Kanagawa (JP)
- Yasuno, Takayuki  
Yokohama-shi, Kanagawa (JP)

(74) Representative: Hoffmann, Eckart, Dipl.-Ing.  
D-82166 Gräfelfing (DE)

### (54) Method and apparatus for reconstructing three-dimensional objects

(57) One of images of an object (19) captured by at least three cameras (111-11n) is defined as a basic image, and a basic backprojection line is defined which passes through a feature point in the basic image corresponding to a noted three-dimensional feature point on the object and the optical center of the camera having captured the basic image. Reference backprojection lines are calculated respective feature points on an epipolar line, obtained by projecting the basic backprojection line onto the image by each of the other cameras, and the optical center of said each camera. The coordinates of intersection points of the basic backprojection line and the reference backprojection lines are calculated and the number of the reference backprojection lines intersecting at each intersection point on the basic backprojection line is counted. The point having the maximum number of intersections is determined to be the position of the noted three-dimensional feature point. Similar processing is performed for all of the feature points in the basic image, by which the positions of corresponding three-dimensional feature points of the object are determined.

FIG. 1



EP 0 631 250 A3





European Patent  
Office

# EUROPEAN SEARCH REPORT

Application Number  
EP 94 10 9510

| DOCUMENTS CONSIDERED TO BE RELEVANT  |  |   |   |
|--|--|---|---|
| Category   | Citation of document with indication, where appropriate, of relevant passages  | Relevant to claim                                   | CLASSIFICATION OF THE APPLICATION (Int. CL.5)         |
| A  | IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE,<br>vol. 13, no. 1, January 1991 NEW YORK US,<br>pages 73-85, XP 000205319<br>NICOLAS AYACHE ET AL. 'Trinocular Stereo Vision for Robotics'<br>* page 77, paragraph VII - page 81, paragraph X *  | 1,2,14, 15  | G06F15/72<br>G06F15/70                                |
| A  | PROCEEDINGS OF THE 1990 INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION, MAY 13-18, 1990 CINCINNATI (US); IEEE COMPUTER SOCIETY PRESS, LOS ALAMITOS (US),<br>pages 2045-2050, XP 000143849<br>UMESH R. DHOND ET AL. 'Binocular versus Trinocular Stereo'<br>* page 2046, paragraph B *                             | 1,2,14, 15  |   |
| A  | PROCEEDINGS OF THE THIRD WORKSHOP ON COMPUTER VISION: REPRESENTATION AND CONTROL, OCTOBER 13-16, 1985, BELLAIRE (US); IEEE COMPUTER SOCIETY PRESS, WASHINGTON (US),<br>pages 168-178,<br>ROBER C. BOLLES ET AL. 'Epipolar-Plane image analysis: A technique for analysing motion sequences '<br>* the whole document * | 1,2,14, 15  | TECHNICAL FIELDS SEARCHED (Int. Cl.5)<br>G06F<br>G06T |
| A  | SYSTEMS & COMPUTERS IN JAPAN, vol. 19, no. 12, December 1988 NEW YORK US,<br>pages 13-22, XP 000112795<br>YOSHIFUMI KITAMURA ET AL. 'Trinocular Vision Using edge continuity and local features'<br>* page 13, paragraph 2 - page 15, paragraph 2.3 *  | 1,2,14, 15  |   |
| The present search report has been drawn up for all claims   |  |   |   |
| Place of search<br>THE HAGUE   |  | Date of completion of the search<br>8 November 1995 | Examiner<br>Chateau, J-P                              |
| <p><b>CATEGORY OF CITED DOCUMENTS</b></p> <p>X : particularly relevant if taken alone<br/>Y : particularly relevant if combined with another document of the same category<br/>A : technological background<br/>O : non-written disclosure<br/>P : intermediate document</p> <p>T : theory or principle underlying the invention<br/>E : earlier patent document, but published on, or after the filing date<br/>D : document cited in the application<br/>L : document cited for other reasons<br/>&amp; : member of the same patent family, corresponding document</p> |  |   |   |

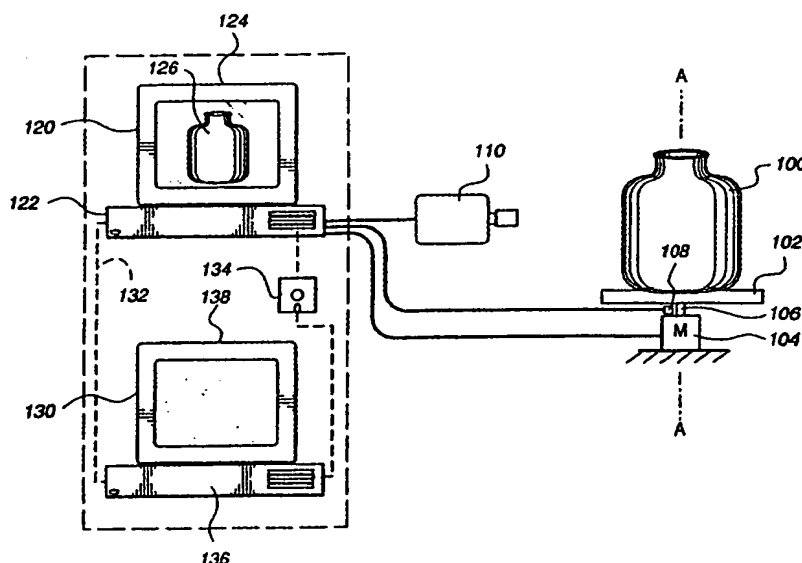
EPO FORM 150 (3.92) (PM/CM)



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

|   |           |   |
|---|-----------|---|
| (51) International Patent Classification <sup>6</sup> :<br><b>G01B 11/00, G06T 7/00</b>   | <b>A1</b> | (11) International Publication Number: <b>WO 98/17970</b><br>(43) International Publication Date: <b>30 April 1998 (30.04.98)</b>   |
| <p>(21) International Application Number: <b>PCT/US97/19331</b></p> <p>(22) International Filing Date: <b>24 October 1997 (24.10.97)</b></p> <p>(30) Priority Data:<br/><b>08/738,437</b>      <b>25 October 1996 (25.10.96)</b>      <b>US</b></p> <p>(71) Applicant: <b>WAVEWORX, INC. [US/US]; 2431 Mar East Street, Tiburon, CA 94920 (US).</b></p> <p>(72) Inventors: <b>MIRAMONTI, John, L.; 9 Lash Road, West Lebanon, NH 03784 (US). MUELLER, Frederick, E.; 908 Guerrero, San Francisco, CA 94110 (US).</b></p> <p>(74) Agent: <b>VAN PELT, Lee; Hickman Beyer &amp; Weaver, P.O. Box 61059, Palo Alto, CA 94306 (US).</b></p> |           | <p>(81) Designated States: <b>AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).</b></p> <p><b>Published</b><br/><i>With international search report.</i><br/><i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i></p> |

(54) Title: METHOD AND APPARATUS FOR THREE-DIMENSIONAL COLOR SCANNING



## (57) Abstract

The invention provides a three-dimensional digital scanner which includes a multiple view detector which is responsive to a broad spectrum of visible light. The multiple view detector is operative to develop a plurality of images of a three-dimensional object which is being scanned. The plurality of images are taken from a plurality of relative angles with respect to the object, and the plurality of images depict a plurality of surface portions of the object. A digital processor including a computational unit is coupled to the detector and is responsive to the plurality of images so that it develops 3-D coordinate positions and related image information for the plurality of surface portions of the object. A three-dimensional image of the object to be scanned is thus developed by the digital processor. The data developed includes both shape and surface image color information.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

|    |                          |    |  |    |  |    |                          |
|----|--------------------------|----|--|----|--|----|--------------------------|
| AL | Albania                  | ES | Spain                                    | LS | Lesotho                                      | SI | Slovenia                 |
| AM | Armenia                  | FI | Finland                                  | LT | Lithuania                                    | SK | Slovakia                 |
| AT | Austria                  | FR | France                                   | LU | Luxembourg                                   | SN | Senegal                  |
| AU | Australia                | GA | Gabon                                    | LV | Latvia                                       | SZ | Swaziland                |
| AZ | Azerbaijan               | GB | United Kingdom                           | MC | Monaco                                       | TD | Chad                     |
| BA | Bosnia and Herzegovina   | GE | Georgia                                  | MD | Republic of Moldova                          | TG | Togo                     |
| BB | Barbados                 | GH | Ghana                                    | MG | Madagascar                                   | TJ | Tajikistan               |
| BE | Belgium                  | GN | Guinea                                   | MK | The former Yugoslav<br>Republic of Macedonia | TM | Turkmenistan             |
| BF | Burkina Faso             | GR | Greece                                   | ML | Mali   | TR | Turkey                   |
| BG | Bulgaria                 | HU | Hungary                                  | MN | Mongolia                                     | TT | Trinidad and Tobago      |
| BJ | Benin                    | IE | Ireland                                  | MR | Mauritania                                   | UA | Ukraine                  |
| BR | Brazil                   | IL | Israel                                   | MW | Malawi                                       | UG | Uganda                   |
| BY | Belarus                  | IS | Iceland                                  | MX | Mexico                                       | US | United States of America |
| CA | Canada                   | IT | Italy                                    | NE | Niger  | UZ | Uzbekistan               |
| CF | Central African Republic | JP | Japan                                    | NL | Netherlands                                  | VN | Viet Nam                 |
| CG | Congo                    | KE | Kenya                                    | NO | Norway                                       | YU | Yugoslavia               |
| CH | Switzerland              | KG | Kyrgyzstan                               | NZ | New Zealand                                  | ZW | Zimbabwe                 |
| CI | Côte d'Ivoire            | KP | Democratic People's<br>Republic of Korea | PL | Poland                                       |    |                          |
| CM | Cameroon                 | KR | Republic of Korea                        | PT | Portugal                                     |    |                          |
| CN | China                    | KZ | Kazakstan                                | RO | Romania                                      |    |                          |
| CU | Cuba                     | LC | Saint Lucia                              | RU | Russian Federation                           |    |                          |
| CZ | Czech Republic           | LI | Liechtenstein                            | SD | Sudan  |    |                          |
| DE | Germany                  | LK | Sri Lanka                                | SE | Sweden                                       |    |                          |
| DK | Denmark                  | LR | Liberia                                  | SG | Singapore                                    |    |                          |
| EE | Estonia                  |    |  |    |  |    |                          |

Method and Apparatus for Three-Dimensional Color ScanningTechnical Field

5 This invention relates generally to optical scanners, and more particularly to optical scanners for providing a digital representation of three dimensional objects.

Background Art

10 Methods for successfully obtaining two dimensional ("2-D") color image data for objects have been developed. This process is commonly known as two dimensional scanning or digitizing. When an object is scanned, a digital data file is created which contains image data including color information which is associated with a set of two dimensional points or coordinates. The color information is obtained by an optical detector or set of optical detectors that are typically organized in a one or two dimensional array.

15 Matching the color information with the correct two dimensional point or location is not a significant problem in two dimensional scanning since the two dimensional point on which the optical detector is focused is the same point that is associated with the color information obtained by the detector. The color information is mislocated only to the extent that there is some error in the location of the point on which the detector is focused (e.g. an error introduced by the optical  
20 system) and that error can readily be minimized.

The problem of associating color information with three dimensional ("3-D") objects is not so easily solved. This is because prior art methods obtain color information with a two dimensional scanning method, while position information is obtained by a three dimensional scanning method. The mapping of the 2-D color information to the 3-D position information is a  
25 complicated process which is prone to significant error.

Many methods exist for obtaining the three dimensional location of the surface points of the object. One such method is a system which uses a laser range finder to scan the object and record the distance between the known three dimensional location of the range finder and the measured location of the surface of the object. The result of using this method or other methods

of generating three dimensional surface models is a set of three dimensional points which accurately represent the surface of the object. A characteristic of this method and other methods of obtaining a three dimensional surface model is that it is inherently monochromatic, that is, no color information is obtained in the process. If three dimensional color information is desired, then it must be generated by somehow combining or conformally mapping the two dimensional color information onto the three dimensional surface model.

The problem of conformally mapping the two dimensional color information onto the three dimensional surface model is difficult and it is common for mismatching of color information with the three dimensional points to occur. The problem may be visualized by imagining a white statue or bust of a person's head and a color photograph of the same person's face. The photograph cannot simply be projected onto the bust to transfer the correct color information to the correct points on the bust or significant distortion will occur. A significant amount of judgment must be exercised in order to correctly associate the color information from the photograph with the correct surface points on the bust. Similarly, it is difficult to accurately associate color information obtained from two dimensional optical detectors with the correct points on a three dimensional surface model. Another problem in the prior art is that color information is not used to determine surface locations, which means less than the total amount of information that is available is being used. Furthermore, both a 2-D and 3-D system is required, which adds cost.

What is needed is a way of generating a set of three dimensional points representing a surface in such way that the three dimensional points are already associated with color data so that conformally mapping separately generated color data onto the set of three dimensional surface points is not necessary. Furthermore, it is desirable to utilize all available frequencies of light to determine surface point positions to maximize the accuracy of the scanning process and to eliminate a separate 3-D scanning step.

Disclosure of the Invention

Accordingly, the present invention provides a system and method for using the color information from a series of two dimensional color images to derive the three dimensional location in space of the surface points which produced the color images. Because the color information itself is used to derive the three dimensional location of the surface points, there is no need to conformally map separately generated color information onto the derived three dimensional surface points. The points are derived from color information and so are already associated with the correct color information. Also, the use of the color information increases the accuracy of the three dimensional location of the surface points.

In one embodiment, the present invention provides a three dimensional digital scanner which includes a multiple view detector which is responsive to a broad spectrum of visible light. The multiple view detector is operative to develop a plurality of images of a three dimensional object which is being scanned. The plurality of images are taken from a plurality of relative angles with respect to the object, and the plurality of images depict a plurality of surface portions of the object. A digital processor including a computational unit is coupled to the detector and is responsive to the plurality of images so that it develops 3-D coordinate positions and related image information for the plurality of surface portions of the object. A three dimensional image of the object to be scanned is thus developed by the digital processor. The data developed includes both shape and surface image color information.

In another embodiment, a three dimensional color digital scanner includes a color detector responsive to a broad spectrum of visible light to develop a plurality of images of a three dimensional object. A rotary object support having an axis of rotation allows the detector to develop a plurality of images of a three dimensional object. The plurality of images depict a plurality of surface portions of the object. A digital computer is coupled to the detector. The computer tracks patches of the surface portions of the object to determine coordinates of the patches as a function of the rotation of the rotary object support and determines radii of the patches from the axis of rotation.

In another embodiment, a method for scanning a three dimensional object includes developing a plurality of images of a three dimensional object taken from a plurality of relative angles with respect to the object. The plurality of images depict a plurality of surface portions of the object to be scanned. 3-D coordinate positions and related image information about the  
5 plurality of surface portions of the object is computed from the plurality of images such that a three dimensional image of the object is developed that includes both shape and surface image information.

In another embodiment, a method for determining three dimensional coordinates of a surface portion of an object includes obtaining a plurality of images of the surface portion of the  
10 object and identifying a trackable patch of the surface portion in an initial image. An initial set of two dimensional coordinates of the trackable patch in the initial image is determined along with at least one additional set of two dimensional coordinates of the trackable patch in another of the images. A radial coordinate of the trackable patch is determined and then a set of three dimensional coordinates of the trackable patch are determined from the radial coordinate of the  
15 trackable patch.

In another embodiment, a method for determining three dimensional coordinates of a surface portion of an object includes rotating the object about an axis of rotation so that a plurality of images of the surface portion of the object are obtained as the object is rotates about the axis of rotation. A trackable patch is identified and the two dimensional coordinates of the trackable  
20 patch are determined. The movement of the trackable patch is tracked as a function of the rotation of the object. A radial distance of the trackable patch from the axis of rotation is determined based on the movement of the trackable patch as a function of the rotation of the object and three dimensional coordinates of the surface portion of the object are derived from the coordinates of the trackable patch and the radial distance of the trackable patch from the axis of  
25 rotation.

The present inventions provides a system and method for obtaining 3-D surface information that is linked to color information without the need to conformally map 2-D color data

onto a 3-D surface. The accuracy of the system is enhanced by the use of color data and the cost of the system is reduced because the 3-D surface is derived from a series of 2-D images. These and other advantages of the present invention will become apparent upon reading the following detailed descriptions and studying the various figures of the drawings.



Brief Description of the Drawings

Figure 1 illustrates a system for obtaining a series of two dimensional color images of an object and processing those images to obtain a three dimensional model of the surface of the object.

Figure 1A illustrates an alternative embodiment of the present invention which enables the top and bottom portions of an object to be scanned.

Figure 1B illustrates another embodiment of the present invention which produces enhanced shading of an object.

Figure 1C illustrates an arrangement where a detector is translated about a stationary object.

Figure 1D illustrates an embodiment of the present invention which uses a multiple number of detectors instead of moving a single detector.

Figure 2 illustrates in detail an architecture of an image acquisition system.

Figure 3A is a flow diagram illustrating a process of obtaining multiple images of a rotating object.

Figure 3B is a flow diagram illustrating a process for generating three dimensional surface data from the two dimensional images of the object.

Figure 4 is a flow diagram illustrating a process performed on the images before they are stored.

Figure 5A illustrates the vector nature of the color data obtained.

Figure 5B illustrates an example of blue color data at times 0, 1, 2, and 3 for a line of pixels.

Figure 5C illustrates how the data can be compressed by recording only the changes in the color data.

Figure 6 is a flow diagram illustrating a process for identifying the silhouette of the object in each image.

Figure 7 is a flow diagram illustrating a process for finding silhouette edges along each scan line.

Figure 8 is a flow diagram illustrating a process for determining a set of trackable patches.

Figure 8A illustrates how to search an image for trackable patches.

5 Figure 9 is a flow diagram illustrating a process for determining the radius of the location of patches on the surface of the object as the object is rotated.

Figure 9A illustrates a set of patch tracking limits.

Figure 9B illustrates the motion of trackable patches in different images with different angular displacements.

Figure 9C illustrates the determination of an exact position of the patch in an image.

10 Figure 9D is a graph which illustrates the filtering of raw data points.

Figure 9E is a graph which illustrates how the radius is determined from the points representing the path of the trackable patch across angularly displaced images.

Figure 10 is a flow diagram illustrating the post processing that occurs once the radius of the trackable patch is known.

Best Modes for Carrying out the Invention

In Figure 1, an embodiment of the present invention includes a system for obtaining a series of two dimensional color images of an object and processing those images to obtain a three dimensional model of the surface of the object. An object 100 which is to be digitized is placed  
5 on a rotatable platform 102. A motor 104 is provided to drive rotatable platform 102 via a shaft 106. A position encoder 108 detects the angular position of rotatable platform 102 and generates an electrical signal which represents the angular position of rotatable platform 102. An optical detector 110 (e.g. a color video camera) views object 100 and creates a two dimensional color image of object 100.

10 As object 100 is rotated by rotatable platform 102, detector 110 captures a series of color images of object 100. Each color image taken at a different time is associated with an angular rotation of object 100 about an axis of rotation, "A" which runs through shaft 106. Information about the angular position of object 100 is obtained from position encoder 108. Thus, each "snapshot" or image of object 100 taken by detector 110 from a different view is associated with  
15 data about the angle of rotation of object 100 with respect to detector 110. An image input processing system 120 ("computer") controls the image acquisition process and records the acquired images along with the associated angular position data. That is, processing system 120 is connected to detector 110 and receives data for each image or snapshot taken of object 100 from detector 110, and position encoder 108 sends angular position information to processing  
20 system 120, so that processing system 120 can associate the image data from detector 110 with the angular position data taken at the same time. In other embodiments, detector 110 is a film camera and processing system 120 receives data from a digitizer which digitizes the film images from detector 110.

Processing system 120 includes a processing unit 122 and a monitor 124 and also  
25 controls motor 104. A monitor 124 can display a current image 126 being captured by detector 110 or other information about the capturing process.

Once processing system 120 has obtained a series of images, those images are transferred to an image processor 130 ("computer"). Image processor 130 can receive data from processing system 120 in a number of different ways. Image processor 130 can be directly connected to

processing system 120 via direct connection 132, or data from processing system 120 can be transferred to a removable storage medium such as disk 134 which may be read by image processor 130. Processing system 120 may also transfer data to image processor 130 via the Internet or a modem connection. Image processor 130 includes processing unit 136 and also  
5 includes monitor 138.

In other embodiments, processing system 120 and image processor 130 are combined on a single computer. The advantage of separating the functions of processing system 120 and image processor 130 is that the data acquisition and storage function performed by processing system 120 and control of the data acquisition system does not require a complex or powerful  
10 processor. On the other hand, image processor 130 receives data representing a series of two dimensional images and perform complex and computationally intensive operations on that data to produce a three dimensional surface model. Image processor 130 is therefore, given current technology, likely to be a more powerful (and costly ) computer than processing system 120. If that is the case, then it is economically beneficial to utilize a large number of relatively cheap  
15 processors for data acquisition and temporary storage and send data from those relatively cheap systems to a smaller number of image processors which generate the three dimensional surface model from the set of two dimensional color images.

Figure 1A illustrates an alternative embodiment of the present invention which enables the top and bottom portions of an object to be scanned. Again, object 100 is supported by rotatable  
20 platform 102 which is driven by motor 104. In this embodiment, shaft 107 engages the edge of rotatable platform 102, so that motor 104 and shaft 107 do not obscure the image of the bottom of object 100. Rotatable platform 102 is made from a transparent material so that the bottom of object 100 may be viewed through rotatable platform 102. A set of mirrors 109 are placed within the field of view of detector 110 so that images of the top and bottom surfaces of object 100 are  
25 captured by detector 110 in addition to the side views.

Figure 1B illustrates another embodiment of the present invention which is designed to produce contrast enhancing shading of object 100. Again, object 100 is supported by rotatable platform 102 which is driven by a motor 104 via a shaft 106. A second motor 142 also drives a rotatable platform 144 via shaft 146. Encoder 148 generates data representative of the rotational

position of rotatable platform 144 and transmits that data to processing system 120. Likewise, motor 142 receives control commands from processing system 120. A light 150 is mounted on rotatable platform 144 to provide illumination of object 100. Light 150 is oriented to provide contrasting illuminated and shaded portions on object 100 which aid in the tracking of features on the surface of object 100. Because light 150 is mounted on rotatable platform 144 which is separately controllable by processing system 120, different orientations of light 150 with respect to object 100 may be checked to determine which one best enhances the surface features of object 100. When platforms 102 and 144 are rotated in a synchronized manner, the shading remains constant. Additionally, multiple sets of views of object 100 with different shadings can also be obtained by changing the relative position of platforms 102 and 144.

Figures 1, 1A, and 1B each depict embodiments where in the object being imaged is rotated. In another embodiment of the present invention, the object remains stationary and the detector moves around the object. Figure 1C illustrates an arrangement where a detector is translated about a stationary object. It should be noted that as the detector 110 is moved, the optics 111 remain pointed at the object 100. Detector 110 can be move in many ways and object 100 can be supported in many ways. In one embodiment, an unobstructed view of object 100 is obtained by suspending it from very thin wires. Detector 110 is translated about object 100. If object 100 is very large, detector 110 could be mounted on, for example, a helicopter and flown around object 100. It is not necessary that the motion of detector 110 be exactly circular around object 100. The angular and radial components of the motion of detector 110 with respect to object 100 can be computationally analyzed, as will be appreciated by those skilled in the art. As long as the position of detector 110 is measured and recorded, the relative angular position of detector 110 with respect to object 100 can be determined for each image taken by detector 110. Methods of determining the position of detector 110 include using GPS or a laser positioning system. Once the angular component of the motion is analyzed and the radial component is calculated, the system compensates for the radial component and the images generated by detector 110 can be processed similarly to the images generated by a system that includes a rotating object and a stationary detector.

Figure 1D illustrates an embodiment of the present invention which uses a multiple number of detectors instead of moving a single detector. A top view of object 100 is shown and set of detectors 110 are provided at different angular displacements with respect to object 100. The advantage of this embodiment is that no motion is required and the need for motors, encoders, and rotatable supports is limited. The image of object 100 captured by each detector is angularly displaced with respect the images captured by the other detectors and so the images may be processed in a similar manner as successive images taken by one moving detector. The cost of multiple detectors 110 may be less than the cost of a rotatable drive or a mechanism for moving detector 110 and recording the position of detector 110. Another advantage of this approach is that all of the images of object 100 can be created simultaneously.

Figures 1 through 1D depict various embodiments for creating multiple images of object 100 with object 100 and detector 110 at different relative angular displacements. Each of these systems provide two dimensional color images of object 100 observed at different angles. This two dimensional information is converted into a three dimensional surface model of object 100 by the process and apparatus of the present invention.

Figure 2 illustrates in detail the architecture of processing system 120 used in some embodiments. A microprocessor 200 is connected to a memory bus 202 and memory bus 202 is connected to a RAM 204 and a ROM 206. Microprocessor 200 is also connected to an input/output ("I/O") bus 208. A video interface 210 is coupled to I/O bus 208 to control monitor 124, as is detector interface 212. Detector interface 212 buffers and processes data from the detector and also carries output commands to the detector from microprocessor 200. In certain embodiments where a moving detector is used, the detector provides its own control and records its own position. In such embodiments, the detector/processor interface need only be capable of transferring data from the detector, including both image and detector position data, to the processor storage system.

Mass storage 214 (such as a hard disk drive) is also connected to input/output bus 208 and provides storage capacity for the multiple images generated by the optical system. Removable storage 216 (such as a floppy disk drive) also provides a way of transferring data files to and from processing system 120 and another processing system. Alternatively,

communications interface 218 can be used to transfer files as well. Communications interface 218 may be connected to a local area network ("LAN") or wide area network ("WAN") for communication with other workstations. Position controller 220 is connected to input/output bus 208 and provides control to a motor in embodiments where processing system 120 provides control commands for rotating object 100. In such embodiments, position detector 222 receives data from an encoder so that processing system 120 may keep track of the position of object 100. Lighting control 224 is also connected to input/output bus 208 and is used to control the position of lights which may be moved with respect to object 100. Lighting control 224 also controls the intensity of those lights.

The architecture shown for processing system 120 in Figure 2 is capable of supporting any of the embodiments shown in Figures 1-1D. If the object is to be rotated, position controller 220 and position detector 222 provide control of the rotation. Position information about object 100 can be integrated with image data from interface 212 and stored in mass storage 214. Movement and intensity control of the light is controlled by lighting control 224. If an autonomous detector is used, data about the detector position and images captured by the detector can be transferred to processing system 120 via communications interface 218 or removable storage 216. Multiple detector interfaces are provided to control a multiple number of detectors in embodiments which use more than one detector. As described above, a three dimensional surface model can be computed using microprocessor 200 and the data contained in mass storage 214, or, alternatively, the data in mass storage 214 can be transferred to a more powerful image processing system.

Figure 3A is a flow diagram for the process of the present invention of obtaining multiple images of a rotating object. Preferably, the method is implemented on a processing system 120. The process starts at step 300, and the user places the object on the rotatable platform in step 302. The object begins to rotate while it is being imaged by a detector. In step 304, the processor checks whether the required number of images have already been captured or taken. If the required number of images have been captured, then the process is finished at step 306. The two dimensional image data is then ready to be taken to an image processor for generation of a three dimensional surface model. If more images are to be captured, then control is transferred to step

308, and a command is sent to the detector to capture an image. The image is preferably preprocessed and stored in step 310, and control is transferred back to step 304. Once the process depicted in Figure 3A is complete, a set of two dimensional color images of the object has been obtained at different relative angular displacements between the detector and the object.

5        Figure 3B illustrates the process implemented on image processor 130 for generating three dimensional surface data from the two dimensional images of the object developed by the process illustrated in Figure 3A which was implemented on processing system 120. The process begins at step 350 and in a step 352, the processor identifies the silhouette of the object in each of the two dimensional color images of the object provided by the image acquisition system. This  
10        silhouette information will be used to reduce the area in each image that must be processed. In step 354, a set of tracking points is determined within the silhouettes of the object. These tracking points are associated with surface features of the object which are identifiable in different images of the object taken by the optical detector. The portion of the image of the object which corresponds to the identifiable feature is referred to as a trackable patch. Each trackable patch  
15        consists of a group of pixels surrounding a geometric point on the surface which is being scanned. The trackable patch is thus a feature associated with and proximal to the geometric point. The patch is distinguishable from adjacent surface locations and therefore trackable. The trackable patch is tracked through a series of images of the object taken at different relative angular displacements between the object and the detector. Step 356 develops a radius from the  
20        axis of rotation for each trackable patch which will account for the motion of the trackable patch in the images. The radius of the trackable patch is used to approximate the radius to the geometric point associated with the trackable patch.

      Once step 356 is completed, a three dimensional surface consisting of the points which correspond to each trackable patch and interpolations between those points has been generated.  
25        In step 358, coordinates of the points corresponding to the modeled surface are output in a standard format. In some embodiments, the raw trackable patch points are output. In other embodiments, points are interpolated between the raw trackable patch points, so that a regular array of points representing a three dimensional surface is generated. Each trackable patch point is already associated with a color value because the trackable patch points are associated with the



trackable patch color information. The color value of the trackable patch is simply attributed to the point determined for that patch. Step 358 outputs the coordinates and their color values. The process is completed at 360.

5 The process of Figure 3B generates three-dimensional surface data from a set of a two dimensional images in a manner that associates color information with the three dimensional points during the generation of those points. Indeed, the color information is used to identify the points which are tracked thereby enhancing the tracking process. The system does not identify the surface of the object independent of color image information and then match the surface points to color information. In the system depicted in Figure 3B, optical detectors obtain color images  
10 of the object. The 3-D surface points of the object are not determined independently from that color image information. In fact, the surface points of the object are not directly determined, rather, the locations of the features located on the surface are determined. Those surface features include color information. Therefore, the locations of color features are directly determined and do not need to be mapped onto a surface model which is independently generated. The regularly  
15 spaced points of the surface model are then generated from the points representing the color surface features by interpolating between the surface feature points.

Figure 4 illustrates a process 310 that can be performed on the images before they are stored. The process starts at 400. Step 402 analyzes the difference between the current acquired image and the previous acquired image. Step 404 applies a compression technique to the data to  
20 reduce the amount of storage space taken up by the data. The process is finished in step 406. Thus, in this embodiment using process 310, the data is compressed for more compact digital storage.

Figures 5A, 5B, and 5C further illustrate the color image pixel data and a compression technique implemented in one embodiment of the invention. Figure 5A illustrates the vector  
25 nature of the color data obtained. In the example, shown, red-green-blue ("RGB") color data is used. This or any other color scheme such as cyan-magenta-yellow black ("CMYK") can be implemented in accordance with the present invention. The three axes shown in Figure 5A are labeled R, G, and B. Color values can be represented as vectors directed to points in this color space. For example, vector 500 represents the color value of a pixel. The values of R, G, and B

are represented by the point to which vector 500 is directed. Vector 500 is the sum of the green component vector 502, the blue component vector 504, and the red component vector 506. Thus, for each point, there is a vector in color space describing the color of the point. The position or pixel location of the point is stored along with the three dimensions of color information which represent the intensities of the three color signals, RGB.

Figure 5B illustrates an example of what the color data for the color blue looks like at times 0, 1, 2, and 3 for a line of pixels. At time 0, no blue color surface feature is being imaged by the line of pixels. At time 1, a surface feature with a blue component has moved into the position being imaged by the line of pixels, so that 8 blue pixels are observed after X number of no blue pixels are observed. Following the 8 blue pixels, Y pixels of no color are again observed. In this example, only one intensity of blue is shown for the purpose of simplicity. In an actual example, varying intensities of blue could also be observed and recorded. At time 2, the surface feature has moved to the right and now X + 2 pixels are observed with no blue and 8 pixels are observed with blue again. Y - 2 pixels of no color are observed next. At time 3, the blue feature has moved to the right one more pixel and so X + 3 pixels of no color are observed followed by the 8 blue pixels followed by Y-3 pixels of no color.

It can be appreciated that recording the color value of each pixel for each color would generate a large amount of data. Figure 5C illustrates how the data can be compressed by recording only the changes in the color data, and then using run length encoding for those changes. Run length encoding techniques are well known to those skilled in the art. At time 0, no color is observed, and so the data stored is simply no color or black. At time 1, X pixels are still no color, so that number of pixels is recorded and associated with no change. Next, 8 pixels have changed color to blue and so 8 pixels are recorded as having a positive blue change. Finally, the remaining Y pixels have not changed, so Y pixels, no change is recorded. At time 2, the X pixels on the left side which had no color have not changed and so X pixels are recorded as having no change. Since the blue image has shifted two pixels to the right, the next two pixels which were formerly blue now have no color. Both pixels are recorded as two negative blue pixels since the color change for them was negative blue representing the loss of blue color. The next 6 pixels were formerly blue at time 1 and remained blue at time 2 and so no change is

recorded for the next 6 pixels. The next two pixels were formerly not blue and have become blue. Therefore, the next two pixels are recorded as having positive blue value representing a positive blue change between time 2 and time 1. The next Y-2 pixels were blank during time 1, and remain blank during time 2, so no change is recorded for the next Y- 2 pixels.

5           At time 3 the blue feature has moved one more pixel to the right. The first X+2 pixels in the linear array had no color at time 2 and still have no color at time 3. Therefore, no change is recorded for the first X+2 pixels. The next pixel was formerly blue but, since the blue feature has moved one pixel to the right, that pixel is now has no color. Therefore a negative blue color change is recorded for that pixel. The next 7 pixels were blue at time 2 and remain blue at time 3.  
10           Therefore, no color change is recorded for those 7 pixels. The next pixel had no blue color at time 2 and has a blue color component at time 3. A color change of positive blue is therefore recorded for that pixel. Finally, the Y-1 remaining pixels were formerly blank at time 2 and remain blank at time 3 and so no change is recorded for those Y-1 pixels.

          From this example, it can be seen that as a surface feature moves across the pixels at  
15           successive times, it is possible to represent the image of the feature by recording only the changes in color values for each pixel at each successive time. Successive pixels which do not change color can be simply noted without the need to store three dimensions of color value for each individual pixel. This is only one example of a data compression scheme which is used in one embodiment of the present invention. Other compression techniques can be used as alternatives  
20           to or can be used in combination with this compression technique. It is only required that the data be stored in a format which will be accessible and expandable later on for processing of the two dimensional images to determine a three dimensional surface. Alternatively, if storage capacity is not a problem, data compression can be eliminated.

          In other embodiments of the invention, surface attributes other than color are derived and  
25           stored in connection with the three dimensional surface data. For example, specularly, emissivity, and absorption can be derived from the surface geometry and the orientation of the local surface in multiple views at multiple angles with respect to the surface. In certain embodiments, these attributes also aid in distinguishing trackable patches.

Figure 6 illustrates the process 352 that identifies the silhouette of the object in each image. The process starts at 600. Step 602 tests whether the process is finished with the images. If the processor is finished with the images, then the process terminates at step 604. If more images remain to be processed, step 606 applies a filter kernel to the image which is herein  
5 incorporated by reference. As is known to those skilled in the art, a filter kernel is a set of coefficients which is convolved with an image (an XY array of pixels). Certain filter kernels, such as the Canny and Sobel kernels, highlight edges by producing a high correlation at locations which represent edges in images. Edge detection using filter kernels is described in: "On  
10 Detecting Edges" by Nalwi, PAMI, Chapter 8, pp. 699-714; "Finding Lines and Edges in Images" by Canny, MIT Technical Report TR-720, 1984; "On Calibrating Computer Controlled Cameras for Perceiving 3D Scenes" by Sobel Artificial Intelligence, Vol. 5, pp. 185-198, 1974; and "A Comprehensive Approach to Edge Detection" by Canny, PAMI, Chapter 8, pp. 679-698, 1986. By correlating the filter kernel with the image at all points, it is possible to determine the edge of the object for the object silhouette, by analyzing the correlation value. The silhouette  
15 edges are determined in step 608. Control is transferred back to step 602 to check whether there remain other images to be processed for silhouettes.

The purpose of determining the silhouette of the image is to constrain the amount of area over which surface features are searched for and identified for the object. A flying insect or a noise spike recorded by the detector may exist outside the area of the object shown in a given  
20 image. Noise spikes or insects may appear to be a good candidate to be a trackable surface feature, but if the silhouette or edge of the object has been determined and the insect or noise spike is outside of that edge, then it will not be designated as a possible trackable surface feature. This simplifies the search for surface features and reduces computer processing of the image data by limiting the area to be searched and also prevents extraneous surface feature points from being  
25 identified and mapped as surface points. In certain embodiments, the search for surface features is further limited to an area within the silhouette as is described below.

Figure 7 illustrates further detail regarding process 608 of Figure 6 wherein silhouette edges are found along each scan line. The process starts at 700. In step 702, the processor begins at the left edge of the image and correlates the kernel with the pixels of the image to find a

correlation maximum which corresponds to a potential left edge. In step 704, the processor moves in from the right edge of the image and correlates the kernel with the image data to find potential right edges. In step 706, a heuristic algorithm is used to determine left and right edges of the object. In this step, edges are distinguished from bugs or blips by the nature of the edge.

- 5 In one embodiment this is accomplished by distinguishing between the continuous nature of an edge verses the isolated spatial nature of a spec or blip. Any alternative heuristic algorithm which exists for determining edges in such instances may be used within the scope of the present invention.

Figure 8 illustrates the process 354 of Figure 3B implemented to determine a set of  
10 trackable patches. The process starts at 800. Step 802 determines whether the processor is finished with all images. If the processor is finished, the process ends at step 804. If the processor is not finished, control is transferred to step 806 and the processor locates the vertical center line of the image being processed. Next, the processor steps through the image to look for trackable patches starting at the center of the image and searching outward.

- 15 The size and dimensional nature of the trackable patches varies in different embodiments. In certain embodiments, trackable patches are searched for in individual scan lines, so that the trackable patches are one dimensional. In other embodiments, the trackable patches are searched for across several scan lines at once so that the patches are two dimensional. The size of the trackable patch searched for is fixed in some embodiments at a certain pixel area, for example  
20 10x10. In other embodiments, the area of the trackable patches is not predetermined and may vary in size for different areas of the image. This enables the system to be flexible and select trackable patches of an optimal size for the type of trackable features found on the object. In general, when an "area" is referred to as being checked for trackable patches, any size or dimension area is intended, including a line of pixels.

- 25 Step 808 checks whether the processor is finished testing all parts of the image. If the processor is finished, then control is transferred back to step 802. If the processor is not finished, then the next unsearched area of the image is searched and analyzed to determine if it contains any trackable patches in step 810.

A number of ways exist for determining whether an area or section of the image is a good candidate for a trackable patch. A good trackable patch is a series of pixels which have RGB vector values which are readily identifiable and distinguishable from the RGB vector values in other areas of the image. Therefore, a distinctive coloration of a patch is an asset to its trackability. The trackable patch RGB vector values will be correlated with the RGB values in other images which represent other views of the object. If the trackable patch is readily distinguishable, then the correlation will yield a distinct spike at the new location of the trackable patch in the other images and the patch will be successfully found or tracked.

A number of methods exist for determining how unique or identifiable a given area in an image is, and therefore whether the area is a good candidate to be designated as a trackable patch. One method is to compute the variance values of the pixels within the area. A high variance value will indicate that the pixel values are changing and thus will at least be readily distinguishable from a solid or unchanging color background. A more sophisticated method is to convolve different kernels with the image area being tested to judge the quality of the surface features and its likelihood of being trackable. Standard kernels such as those defined in Canny and Sobel, supra, can be defined as a standard kernel used to check for trackable patches. Alternatively, a learning or smart system in some embodiments has the capability of developing customized kernels on the fly which are optimized for the type of surface features found on the object being scanned. The size of trackable patches, and the kernels used to identify them may be held constant or may be varied.

As noted above, step 806 locates the vertical center line of the image. The system begins at the vertical center line of the image and the search proceeds outward for trackable patches. The reason that trackable patches are first looked for at the center line of the image is that surface features are most likely to be completely and accurately sensed without distortion when they are observed head-on by the detector. However, areas to the left or right of the centerline can also be used. As surface features rotate away from the center line of the object, they will eventually pass out of the view of the detector. Before passing out of the view of the detector, the surface features will begin to become partially obscured and distorted as they are viewed obliquely by the detector. For example, at a 45 degree angle, a circle on the surface of the object will appear to the

detector to be an oval. Obscured and distorted surface features are more difficult to accurately track because they change between images and it is also difficult to define a consistent location point for them.

In step 812, the processor marks the trackable patches found in the area searched. This  
5 includes storing the color data for the pixels in the patch and storing the location of the patch. In step 814, the processor increments a distance away from the center line to define a new area to search and control is transferred to step 808. As noted above, step 808 determines if the image is finished being tested. In one embodiment, this is done by checking whether the entire area within a maximum distance from the center has been searched. The distance from the center line over  
10 which the processor will search for trackable patches can also be varied. Figure 8A illustrates an image being searched for trackable patches by the process 810 of Figure 8. The top portion of Figure 8 illustrates object 100 and area 850 which is magnified below. Area 850 consists of a number of areas or sections 852 which, in turn, consists of pixels 854. Area 850 is located at the center line of object 100 and so will be tested by the processor on its first pass. Each section 852  
15 of area 850 will be checked to determine whether it is a good candidate for a trackable patch. The vector color data from the pixels 854 is checked to see if it contains characteristics which are expected to be readily identifiable and distinguishable from other areas.

As noted above, the processor begins at the center line,  $C_i$  and checks outward for good trackable patches because features are most likely to be identifiable at the center. In other  
20 embodiments, the computer can check for patches starting with a location other than the center of the image. In Figure 8A, the trackable patches are shown as squares which have a predetermined area or number of pixels, in this case, five by five. In other embodiments, the trackable patches have variable or irregular sized areas. Different predefined sized areas can also be used. The important feature is that patches on the surface are identified in the image so that the points  
25 corresponding to those patches can be determined and tracked in other images.

Figure 9 illustrates the process 356 of Figure 3 implemented to determine the radius of the location of patches on the surface of the object as the object is rotating. As noted before, a set of two dimensional images of the object are acquired as the object rotates. These images include surface features of the object located in a two-dimensional space and it is desired to derive the

three-dimensional location of the features from the two-dimensional information in the images. Features which are likely to be trackable are designated as trackable patches.

5 A three dimensional surface is derived by locating each of the patches. This is done by tracking the patch across a number of views which represent angularly rotated images of the object. As the object is rotated (or virtually rotated if multiple detectors or a movable detector are used--whenever the object is described as being, rotated, it should be understood that virtual rotation is also included), features which are located a longer distance from the axis of rotation move a greater distance than features which are located a shorter distance from the axis of rotation. A radius from the axis of rotation of the object is derived for each trackable patch which  
10 will account for the movement of the trackable patch in the different images.

The process starts at 900. Step 910 determines whether the processor is finished with all patches. If the processor is finished with all patches, then control is transferred to step 912 and the process ends. If more patches remain to be processed, then control is transferred to step 913 and a patch is chosen, together with an initial image in which the patch is found. In step 914, the  
15 processor checks whether the patch lies outside of the tracking area within the image. If the patch is not outside the tracking limit, then the processor calculates the expected path of the patch in the image in step 916. If the image is the initial image to be processed, then the expected path will simply be the initial location of the patch in that image.

Next, in step 918, the processor finds the exact position of the patch in the image and  
20 stores that position for that image. In step 920, the processor updates the patch kernel. The reason for updating the patch kernel is that the patch will change as it moves across the image from the center towards the tracking limit. The tracking limit defines the area of the image in which patches will be tracked. This area is bounded in one embodiment across a 90 degree range. Although the tracking limit is set so that features within the limit remain relatively  
25 constant, it is possible near the edges of limit that the image of the patch will change more and more rapidly with angular displacement. The changing of the image of the patch makes it difficult to track the patch and increases the possibility of error.

Although within the region over which the patch is tracked, the patch may change considerably from the center of the region to the edge of the region, it is fortunate that, in many



cases, the patch will change in a relatively continuous manner. Therefore, although the patch when viewed at the edge of the patch tracking region will appear very different than the patch when viewed at the center of the patch tracking region, the patches viewed in areas of the patch tracking region which are close to each other will be similar. To take advantage of this, the kernel  
5 used to search for the patch is changed with each successive image so that the patch searched for in each image is actually the patch which was found in the last image, not the initial patch which was defined in the center image. This makes it much easier to correlate the patch and to successfully track the patch.

Once the patch kernel is updated in step 920, the processor obtains the next image in step  
10 922 and control is transferred back to step 914. When step 914 determines that the image limit has been reached, control is transferred to step 940. In step 940, the data is filtered or smoothed. Filtering and smoothing consists of applying known standard statistical and filtering techniques for cleaning up the data and perhaps rejecting bad data points. The data consists of the positions determined across a plurality of images for each patch. This data is associated with the color  
15 pixel information for each patch. In step 942, a recursive least square ("RLS") algorithm is run to fit the points to a curve. In step 944, the radius from the axis of rotation of the object to the surface of the object at the trackable patch is determined which will account for the motion of the trackable path across the images. Control is then transferred to step 910 which proceeds to the next patch or, if all patches have been processed, control is transferred to step 912 and the  
20 process ends.

At the completion of the process 356, three-dimensional surface information about the object has been obtained. That three-dimensional surface information is integrally linked to color data because the vector color data of the trackable patch was used to find the patch in each successive image so that the motion of the patch could be determined as the object rotated in  
25 different images. The radius of the surface portion containing the trackable patch combined with the two-dimensional pixel information obtained from the optical detector is then be transformed to provide a three-dimensional representation of the point in any desired coordinate system.

Figure 9A illustrates a set of patch tracking limits as used by step 914 of Figure 9. .  
Detector 110 views object 100 straight-on at point 900 on the surface of object 100, i.e. along the

centerline  $C_i$  of the object. As noted above, surface features and trackable patches are viewed with the least amount of distortion at this head-on angle. As the trackable patch is rotated away from point 900, it appears to detector 110 to change shape and becomes distorted. In one embodiment, the judgment is made that it is not worthwhile to attempt to track trackable patches beyond point 902 and point 904, which are located at a 45 degree angular limit from point 900. The tracking limit in each image will be a vertical line corresponding to this angle. In other embodiments, other limits are chosen. Limits can also be learned by the system or determined based on the nature of the data so that no predetermined limits need to be set.

Figure 9B illustrates the motion of trackable patches across different images with different angular displacements as calculated in step 916 of Figure 9. Object 100 in this image is a symmetrically shaped vase so that the silhouette of object 100 remains constant but features on the surface of object 100 move as object 100 is rotated. Object 100 is shown with a trackable patch 950, a trackable patch 952 and a trackable patch 954. Trackable patch 954 is shown located at the exact intersection of the horizontal and vertical center lines of the detector view. When an angularly displaced image of object 100 is analyzed to find trackable patch 954, it is not necessary to search the entire image for the trackable patch in a random manner. Trackable patch 954 is expected to move along a line 956 and so the search for trackable patch 954 in a successive image of object 100 should be executed along line 956. Note that line 956 is straight because trackable patch 954 is located at the center point of the detector view in the first image. Similarly, trackable patch 952 is expected to follow a line 958 in images in which trackable patch 952 is angularly displaced, and a line 959 is the path which trackable patch 950 is expected to follow in angularly displaced images. Line 958 is curved, unlike straight line 956, because line 958 does not pass through the center  $C_d$  of the detector view. Features not located in the center are expected to follow epipolar trajectories as is known to those skilled in the art. Line 959 is even more curved than line 958 because line 959 is further from the center of the detector view than line 958. Thus, it is possible to quickly locate the trackable patches in angularly displaced images because the expected path of the trackable patch is known from the epipolar trajectories as object 100 rotates.

The distance the trackable patch moves along its expected path for a given angular displacement, or the speed of movement of each trackable patch as the object rotates, is not

known. The change in position will vary with the radial distance of the trackable patch from the axis of rotation of the object. This is how the radius from the axis of rotation of the object is determined.

Figure 9C illustrates the determination of an exact position of the patch in an image as determined by step 918 of Figure 9. Trackable patch 960 was observed in a former image and its location is defined by its center point 961. The expected path 962 of center point 961 is known. The processor searches for a good correlation of trackable patch 960 with the image data for patches along expected path 962. Patch 964, represented by a dashed line, does not present a good a correlation as patch 966, which is represented by a solid line. Patch 966 is determined to be the patch and the location of the trackable patch in the new image is determined to be point 968 which represents the location of patch 966.

In this example, the center point of the patch is used to represent the location of the entire patch. This approach works as long as the patch is small. In other embodiments, multiple points are defined for a patch or a weighted average point which is based on the colors of the pixels which make up the patch is defined for the patch. In the illustrated embodiment, the location of the patch in the different images and the motion of the patch in the different images which are angularly displaced is represented by the motion of the center point 961 which represents the center of trackable patch 960 to point 968 which represents the center of the angularly displaced patch 966.

Figure 9D illustrates the filtering of the raw data points as performed in step 940 of Figure 9. Point 970 is shown as being inaccurately derived because it is located away from the expected curve. After the filtering process, point 970 has been moved to point 972 which lies in the expected path. A number of standard filtering techniques can be used. The frequency of the rotation of the object is known and all surface feature motion should occur at a spatial frequency corresponding to that frequency. In one embodiment, a brick wall filter is used to filter the data to remove motion occurring at other frequencies.

In embodiments in which the object is not rotated at a constant frequency, or where views of the object are obtained without rotating the object, (for example, by using multiple cameras) then the spatial frequency of the occurrence of features in images will not correspond to a

frequency of rotation of the object. In such embodiments, the raw data is filtered using a modified Z-transform method. Modified Z-transform methods are used to filter data when the displacement is non-uniform.

Figure 9E illustrates how the radius is determined from the points representing the path of the trackable patch across angularly displaced images and corresponds to steps 940 and 942 of Figure 9. Figure 9E is a plot of curve 980 which represents the first derivative of the distance, or the speed of the patch as it is angularly displaced from the center of the image versus the angular displacement of the patch from the center. When the angular displacement is zero, the displacement of the trackable patch is zero and its linear speed is a maximum value. As the angular displacement increases, the speed will change according to the cosine of the angle of displacement.

The radial distance,  $R$ , of the patch from the axis of rotation of the object can be determined from a graph of the speed of the patch as shown on curve 980. Note that the frequency of this curve is well known so the data can be filtered and the amplitude of the curve can be accurately determined. After fitting the actual rate of displacement of the trackable patch to the cosine curve, the amplitude of the cosine curve is determined. The radius of the trackable patch is derived from the amplitude. Other processing methods are used in different embodiments to derive the radius. In one embodiment, the displacement of the patch, not the velocity is fit to a sine curve. An embodiment can use any processing method available to derive a radius which accounts for the patch motion.

Figure 10 illustrates the post processing process that can occur once the radius of the trackable patch is known. The process begins at step 1000. In step 1010, the radial distance data for the trackable patches is combined with the two-dimensional trackable patch data from the images and converted to a desired coordinate system. In step 1020, the color data associated with the trackable patch is converted to a desired color system for storage. Step 1030 performs interpolation and decimation of the points. In one embodiment, all of the trackable patch points are interpolated so that a regular array of surface points can be defined to model the surface of the object. In step 1040, the points derived in step 1030 and the color values associated with those

points in the form determined by step 1020 are stored in a data structure. The process is completed at 105.

In this process, it is especially important to note that no matching of color data with three-dimensional position data is ever required. The color data for the trackable patches is already associated with the trackable patches which define the surface points of the object, and, in fact was used as an aid in calculating the three dimensional locations of those points. When the trackable patch points are decimated and interpolated, the color values are likewise interpolated. No misregistration of color occurs because the color information is already associated with the patches whose locations are determined to model the surface of the object.

While this invention has been described in terms of several preferred embodiments, there are alterations, permutations, and equivalents which fall within the scope of this invention. It should also be noted that there are many alternative ways of implementing both the process and apparatus of the present invention. It is therefore intended that the following appended claims be interpreted as including all such alterations, permutations, and equivalents as fall within the true spirit and scope of the present invention.

Claims

1. A three dimensional digital scanner comprising:

multiple view detector means responsive to a broad spectrum of visible light, said multiple  
5 view detector means being operative to develop a plurality of images of a three dimensional object  
to be scanned, said plurality of images being taken from a plurality of relative angles with respect  
to said object, said plurality of images depicting a plurality of surface portions of said object to be  
scanned; and

a digital processor including a computational unit, said digital processor being coupled to  
10 said detector means, said digital processor being responsive to said plurality of images and  
operative to develop with said computational unit 3-D coordinate positions and related image  
information of said plurality of surface portions of said object;

whereby a three dimensional image of said object to be scanned can be developed by said  
digital processor that includes both shape and surface image.

15 2. The apparatus of claim 1 wherein said multiple view detector means includes at  
least one detector and means for rotating said object.

3. The apparatus of claim 1 wherein said digital processor is operative to determine  
the boundary of the object in at least one of said plurality of images.

20 4. The apparatus of claim 3 wherein said digital processor is operative to determine  
the boundary of the object in at least one of said plurality of images using a mathematical  
function.

5. The apparatus of claim 4 wherein said mathematical function is a correlation  
function of said image information with a predetermined kernel.

25 6. The apparatus of claim 1 wherein said digital processor is operative to identify a  
trackable surface feature within at least one of said plurality of images.

7. The apparatus of claim 6 wherein said digital processor is operative to determine two-dimensional positions of said trackable surface features within at least two of said plurality of images.

8. The apparatus of claim 7 wherein said digital processor is operative to determine a three dimensional position of said trackable surface feature from a difference between said two dimensional positions.

9. The apparatus of claim 1 wherein said detector means includes a color detector, and wherein said related image information includes color.

10. The apparatus of claim 6 wherein said digital processor is operative to determine a radius from a trackable surface feature to an axis of rotation.

11. A three dimensional color digital scanner comprising:

a color detector responsive to a broad spectrum of visible light to develop a plurality of images of a three dimensional object to be scanned;

a rotary object support having an axis of rotation, said object support allowing said detector to develop a plurality of images of a three dimensional object supported by said object support, where said plurality of images depict a plurality of surface portions of said object;

a digital computer coupled to said detector and responsive to said plurality of images, said digital computer tracking patches of said surface portions of said object to determine coordinates of said patches as a function of the rotation of said rotary object support, said digital computer further determining radii of said patches from said axis of rotation.

12. The three dimensional color digital scanner of claim 11 wherein said digital computer is operative to determine the two dimensional boundary of said object within at least one of said plurality of images.

13. The three dimensional color digital scanner of claim 11 wherein determining the radii of said patches includes determining a first derivative.

14. A method for scanning a three dimensional object comprising:

developing a plurality of images of a three dimensional object to be scanned, said plurality of images being taken from a plurality of relative angles with respect to said object, said plurality of images depicting a plurality of surface portions of said object to be scanned; and

5 computing 3-D coordinate positions and related image information of said plurality of surface portions of said object from said plurality of images such that a three dimensional image of said object to be scanned can be developed that includes both shape and surface image information.

15. The method of claim 14 wherein computing 3-D coordinate positions and related  
10 image information of said plurality of surface portions includes determining a radius of at least one of said surface portions from an axis of rotation.

16. The method of claim 14 wherein computing 3-D coordinate positions and related image information of said plurality of surface portions includes determining a tracking limit for the tracking of said plurality of surface portions.

15 17. The method of claim 14 wherein computing 3-D coordinate positions and related image information of said plurality of surface portions includes correlating a kernel representing an image of at least one surface portion in at least one image with an image of said surface portion in at least one other image.

18. A method for determining three dimensional coordinates of a surface portion of an  
20 object comprising:

obtaining a plurality of images of said surface portion of said object;

identifying a trackable patch of said surface portion in an initial image of said plurality of images;

determining an initial set of two dimensional coordinates of said trackable patch in said  
25 initial image;



analyzing said plurality of images to determine at least one additional set of two dimensional coordinates of said trackable patch in at least one other of said plurality of images;

determining a radial coordinate of said trackable patch based on said sets of two dimensional coordinates; and

5 determining a set of three dimensional coordinates of said trackable patch from said radial coordinate of said trackable patch.

19. The method of claim 18 wherein analyzing said plurality of images to determine at least one additional set of two dimensional coordinates of said trackable patch in at least one other of said plurality of images includes a correlation.

10 20. The method of claim 18 wherein determining a radial coordinate of said trackable patch based on said sets of two dimensional coordinates includes determining a radial speed of said trackable patch across said plurality of images.

21. A method for determining three dimensional coordinates of a surface portion of an object comprising:

15 rotating said object about an axis of rotation;

obtaining a plurality of images of said surface portion of said object as said object is rotated about said axis of rotation;

identifying a trackable patch within said surface portion of said object;

determining the two dimensional coordinates of said trackable patch within said images;

20 tracking the movement of said trackable patch as a function of the rotation of said object;

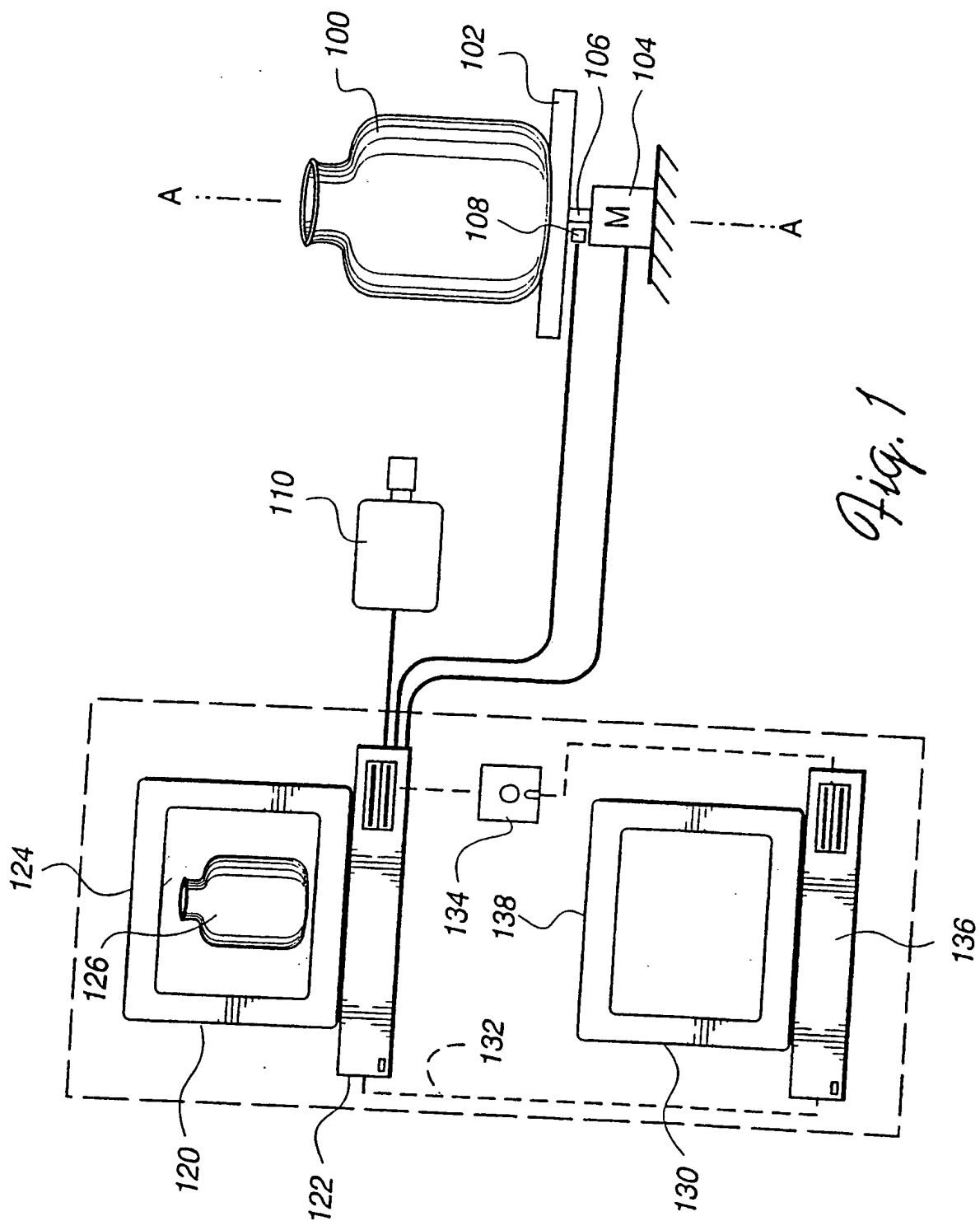
determining a radial distance of said trackable patch from said axis of rotation based on the movement of said trackable patch as a function of said rotation of said object; and

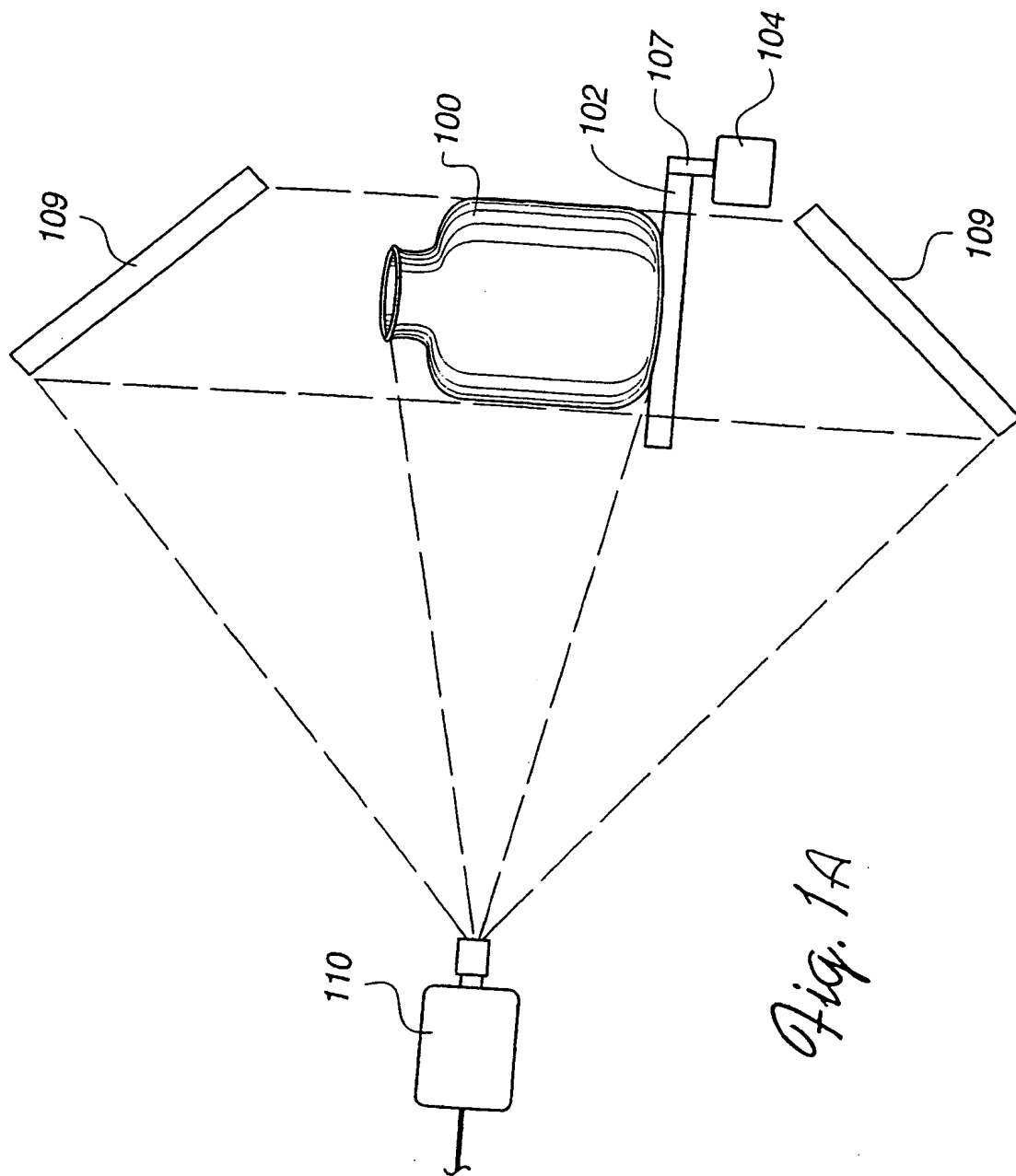
determining said three dimensional coordinates of said surface portion of said object from said coordinates of said trackable patch and said radial distance of said trackable patch from said axis of rotation.

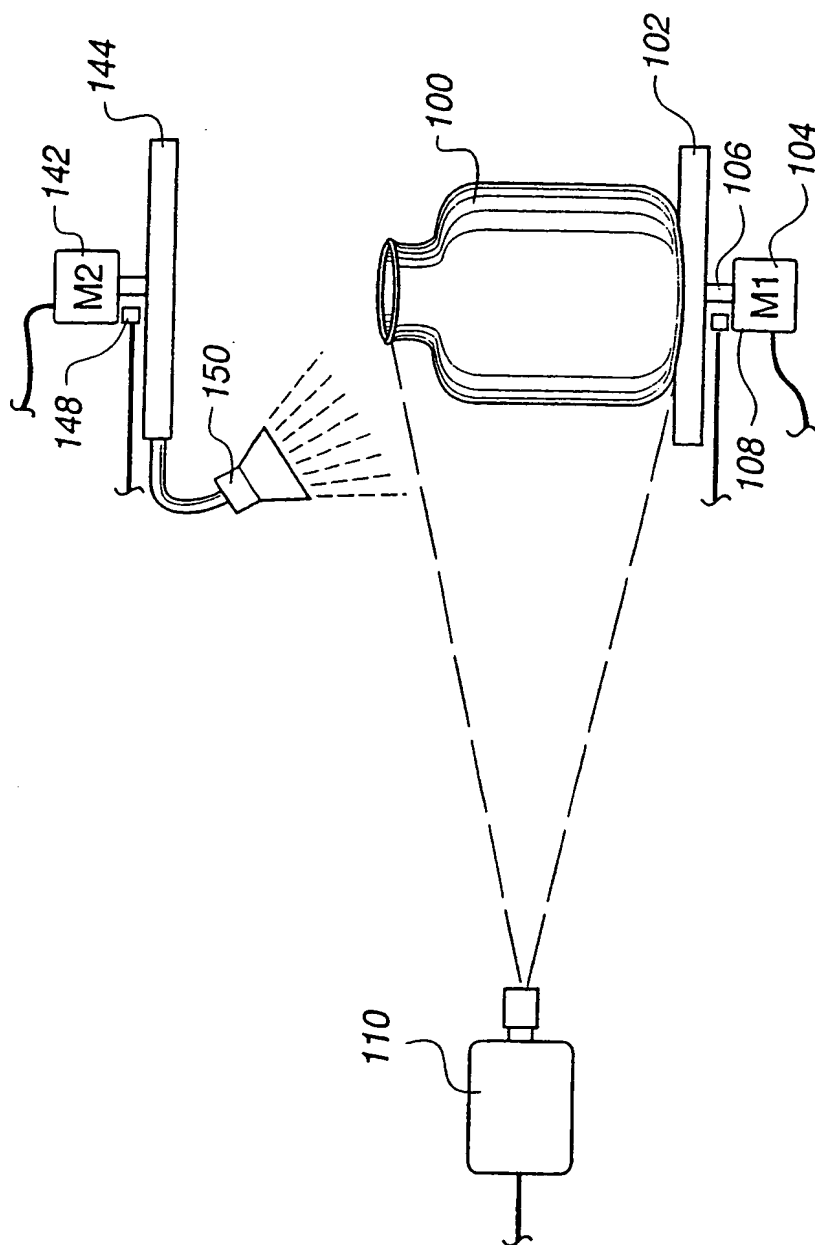
22. The method of claim 21 wherein tracking the movement of said trackable patch as  
5 a function of the rotation of said object includes correlating an image of said trackable patch with portions of an image of said object.

23 The method of claim 21 wherein determining a radial distance of said trackable patch from said axis of rotation based on the movement of said trackable patch includes determining the radial speed of said trackable patch.

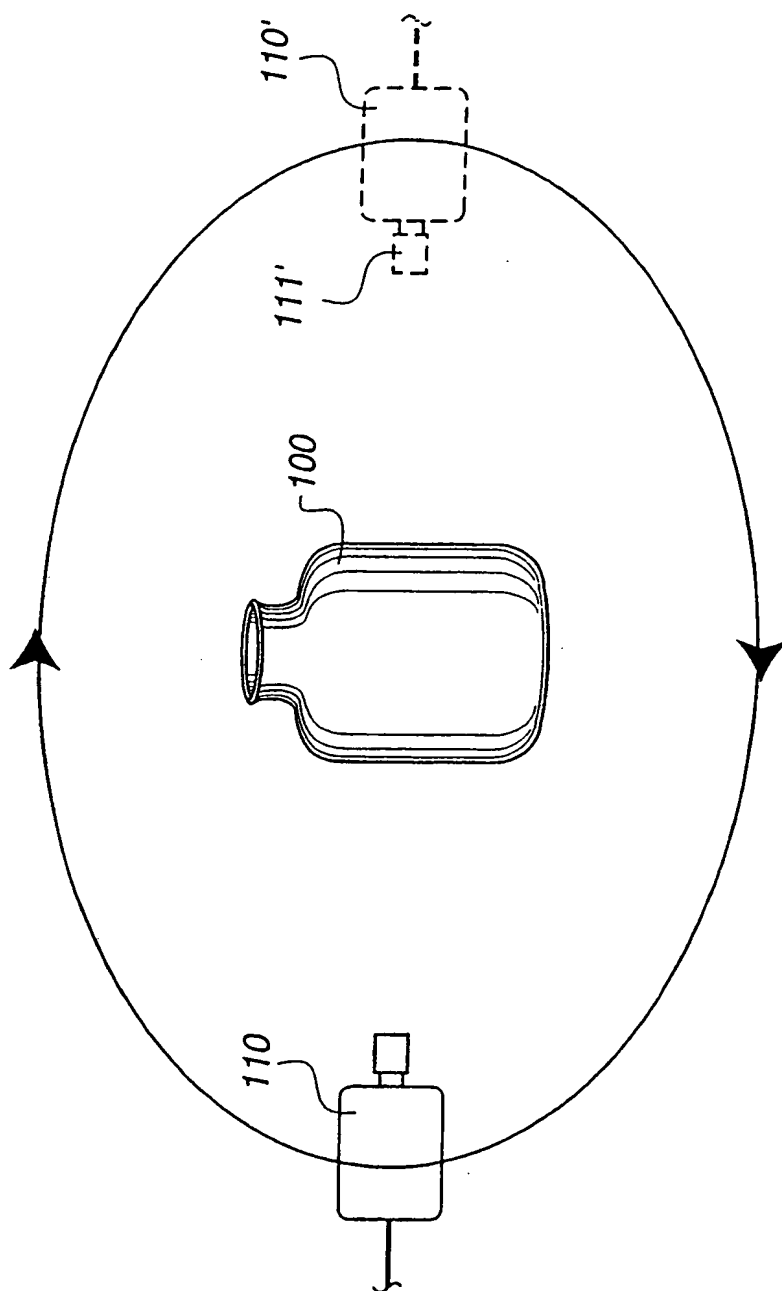
10



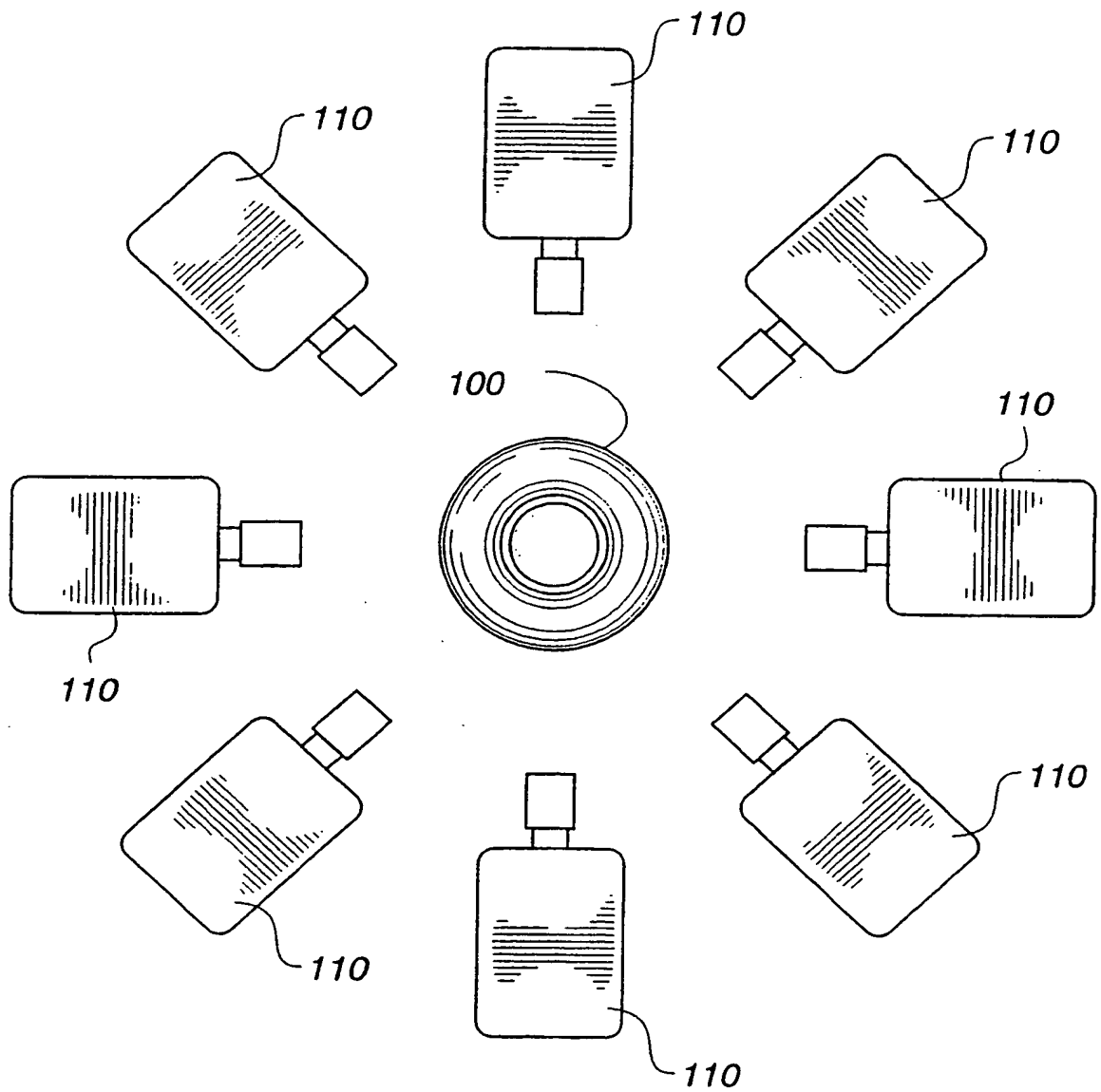
*Fig. 1A*

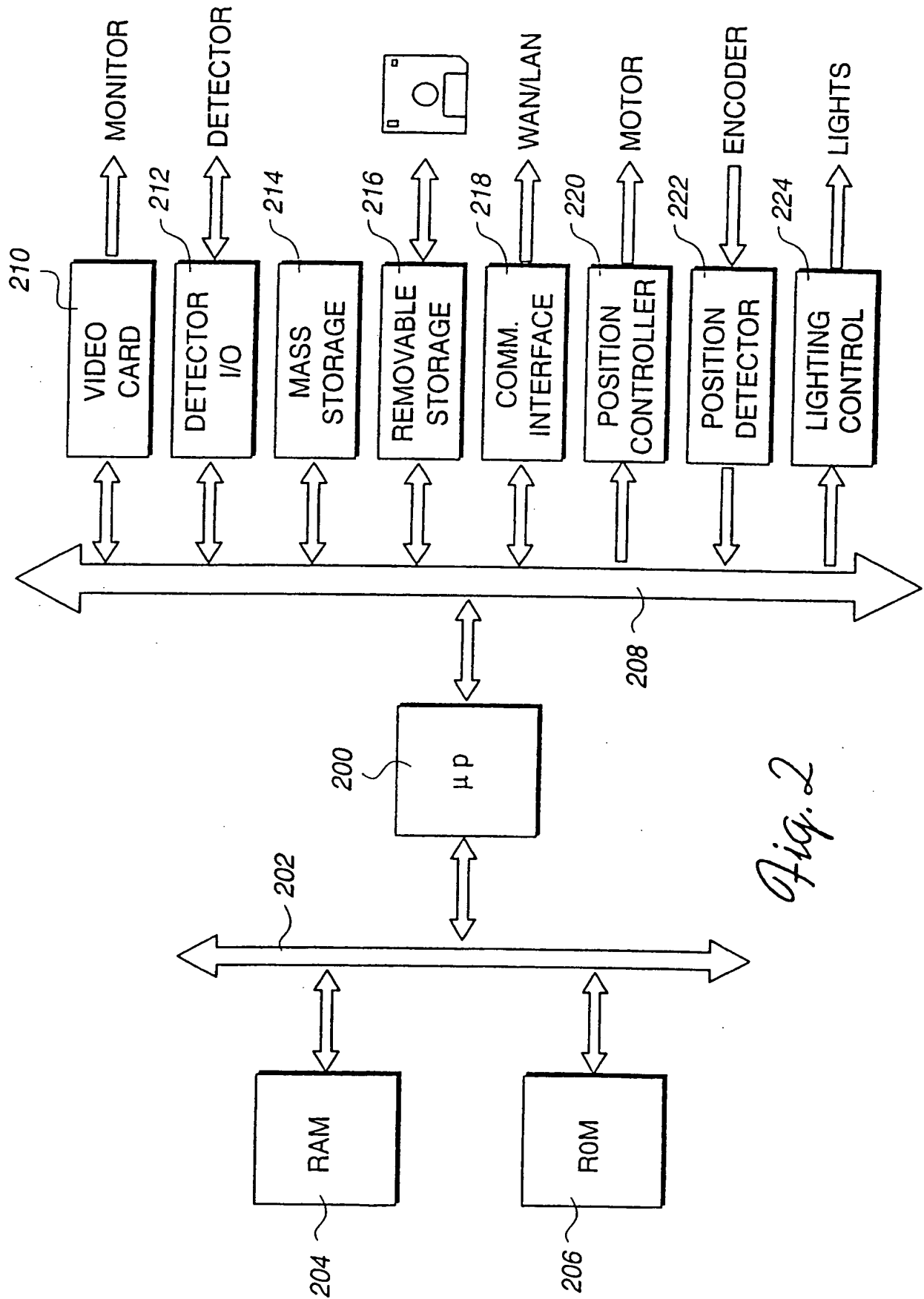


*Fig. 1B*



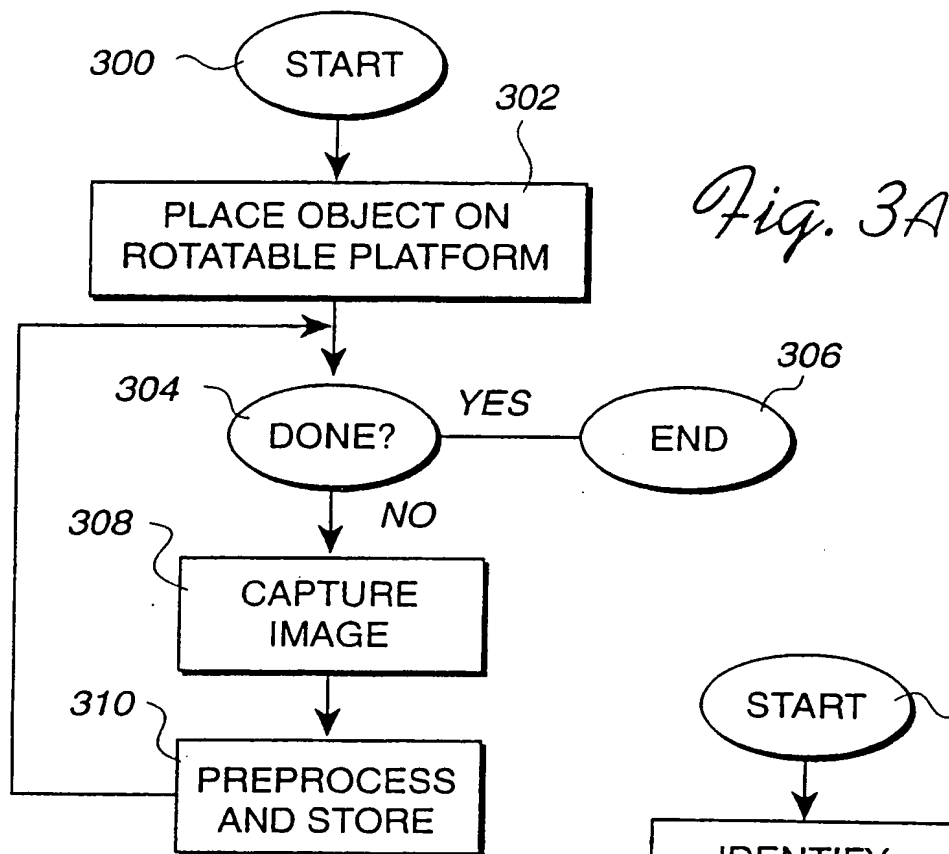
*Fig. 1C*

*Fig. 1D*

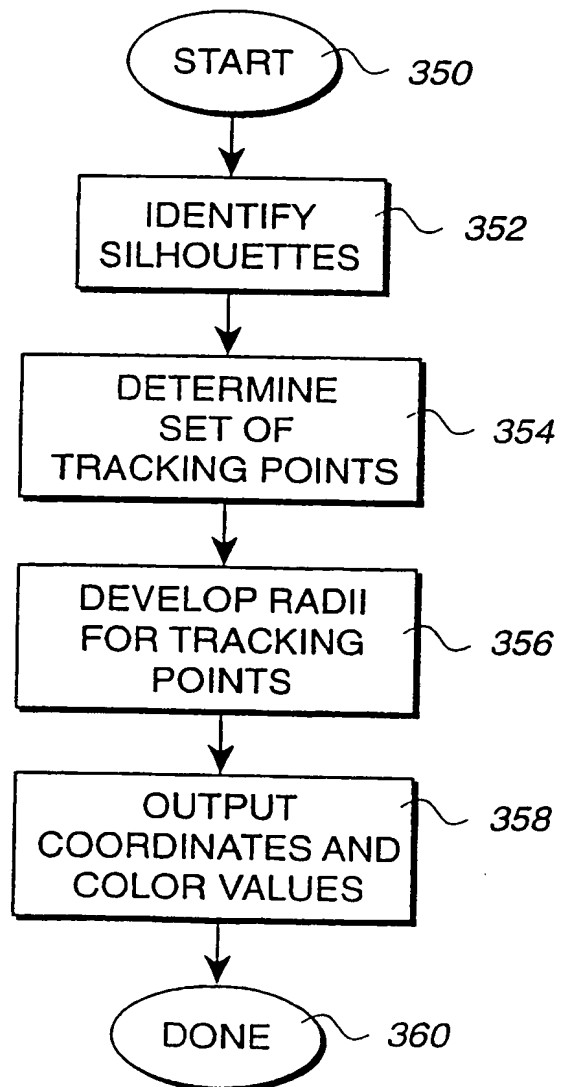


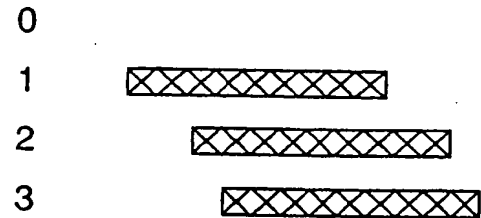
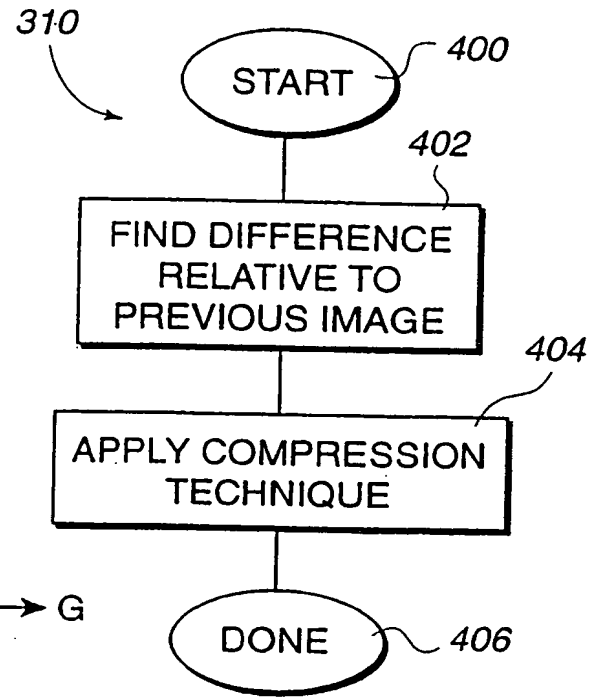
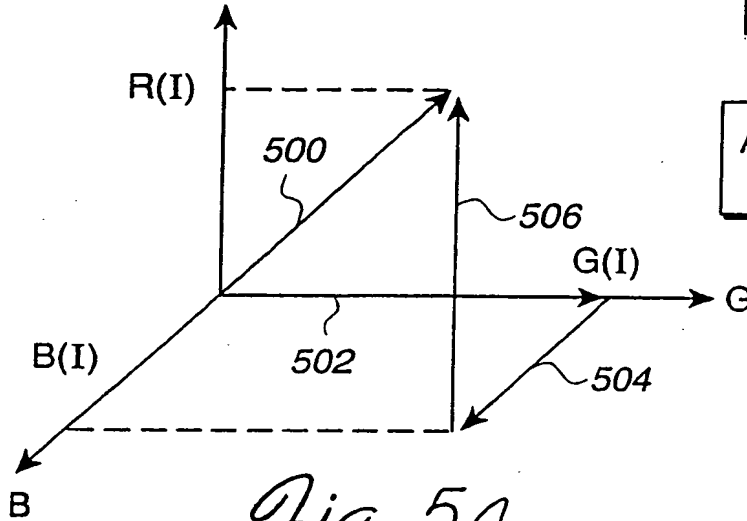
*Fig. 2*





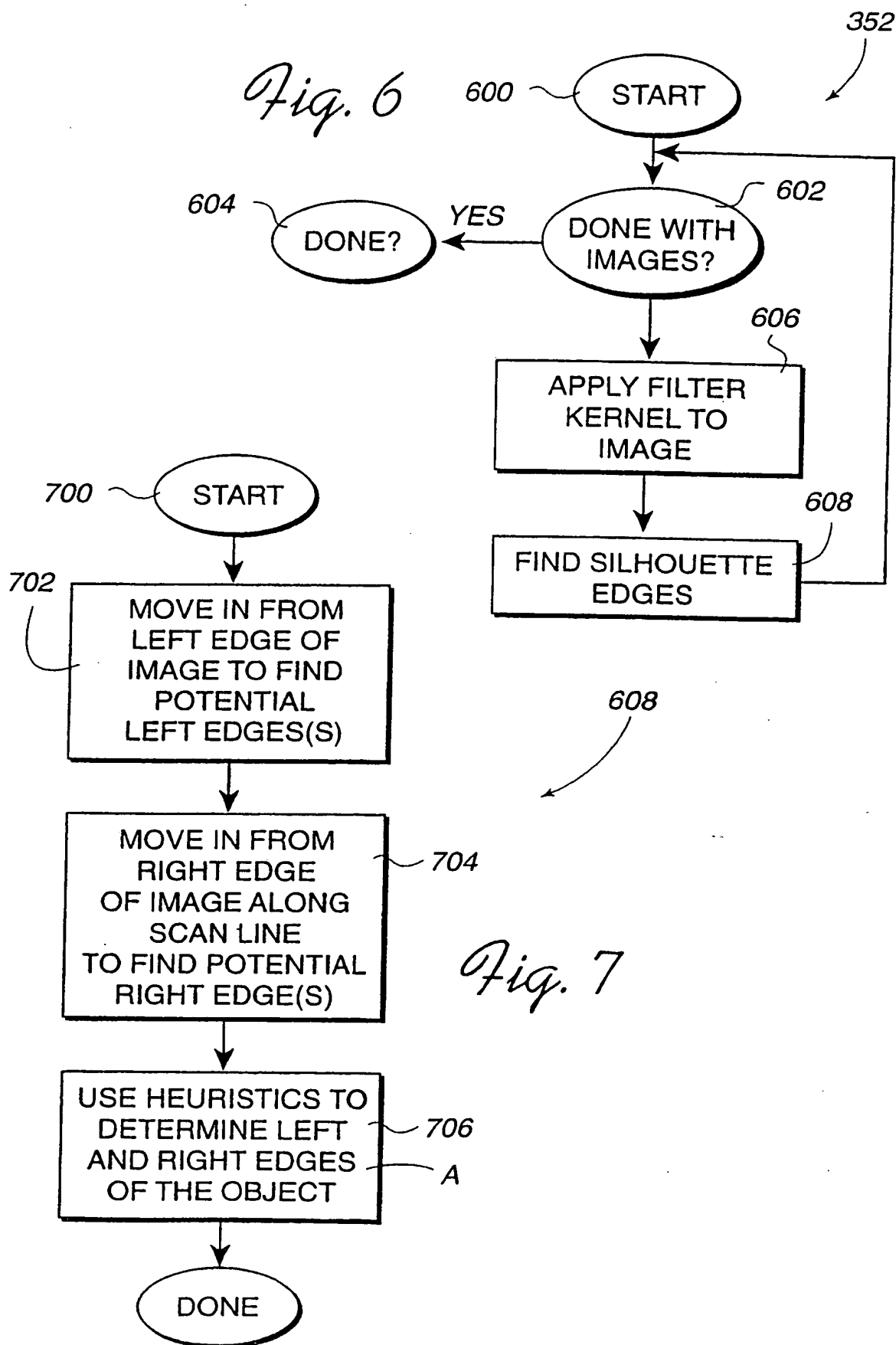
*Fig. 3B*

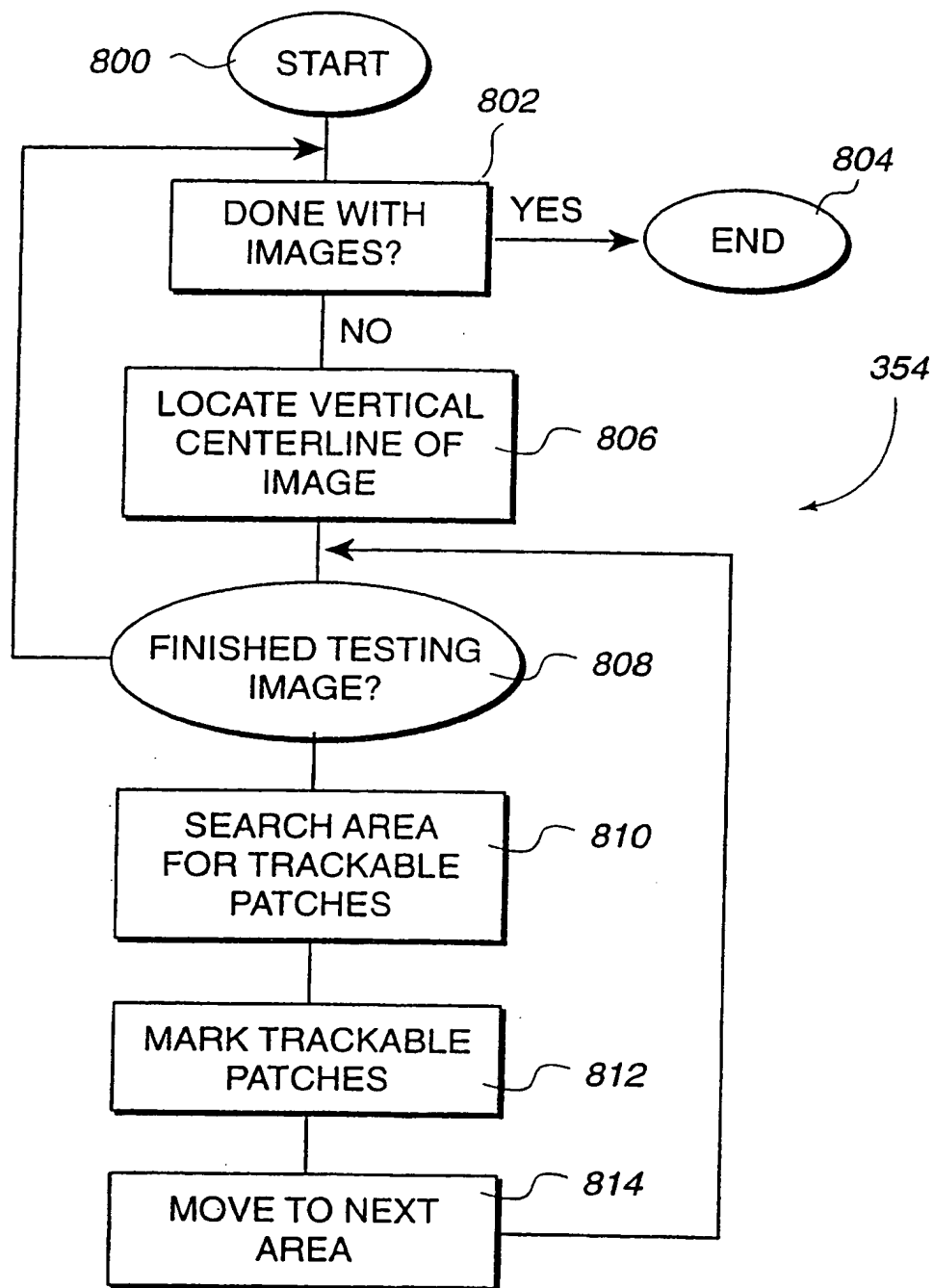


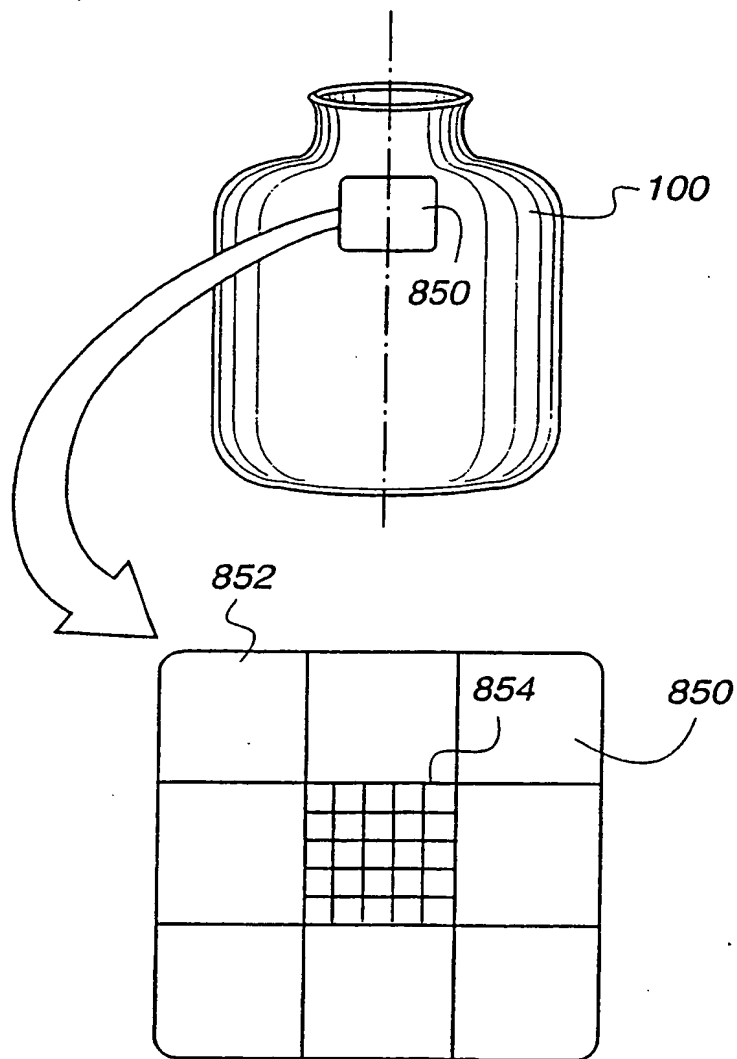


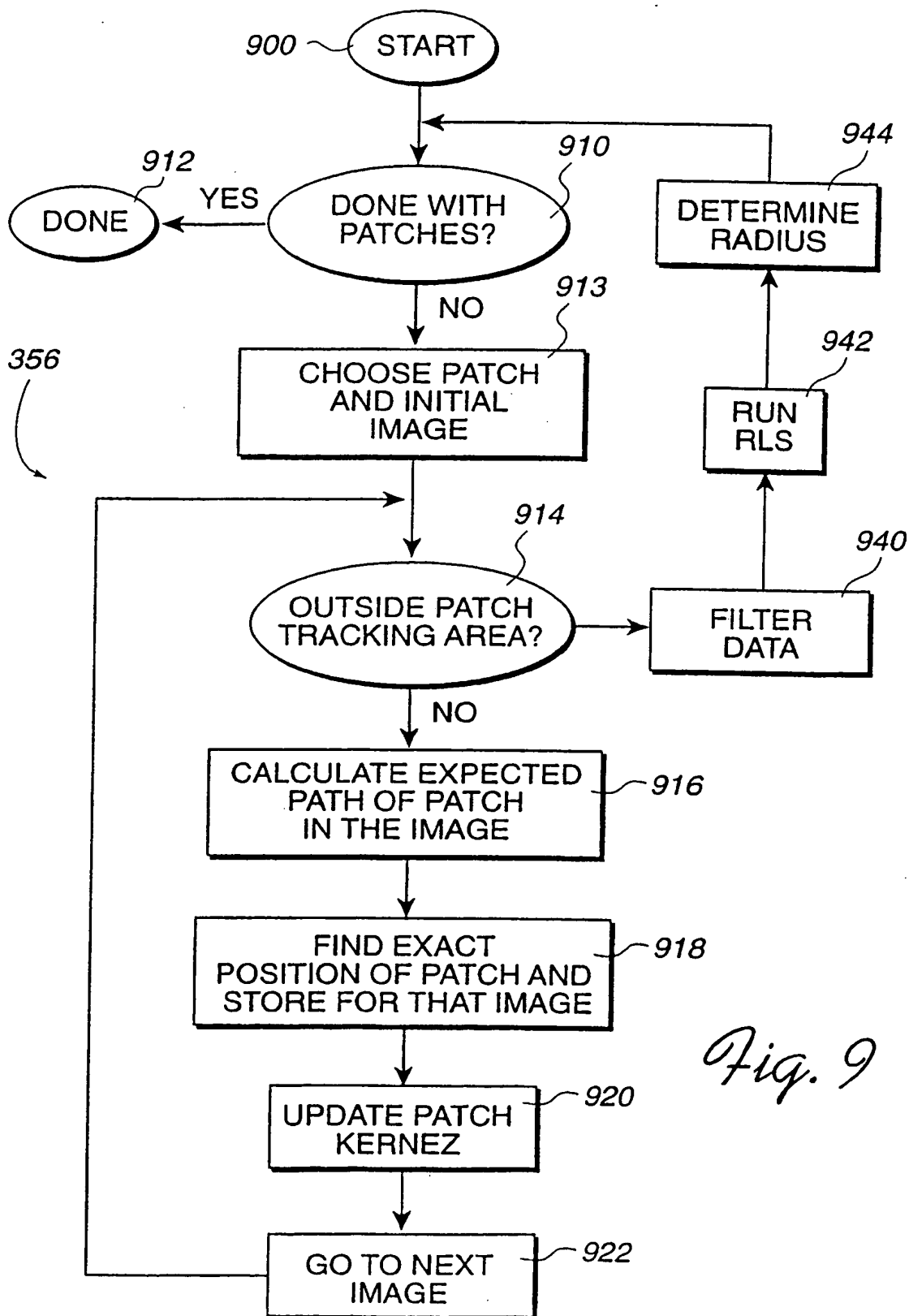
|   |          |       |       |       |         |  |
|---|----------|-------|-------|-------|---------|--|
| 0 | NO BLUE  |       |       |       |         |  |
| 1 | X(NC)    | 8(+B) | Y(NC) |       |         |  |
| 2 | X(NC)    | 2(-B) | 7(NC) | 2(+B) | Y-2(NC) |  |
| 3 | X+2 (NC) | 1(-B) | 7(NC) | 1(+B) | Y-1(NC) |  |

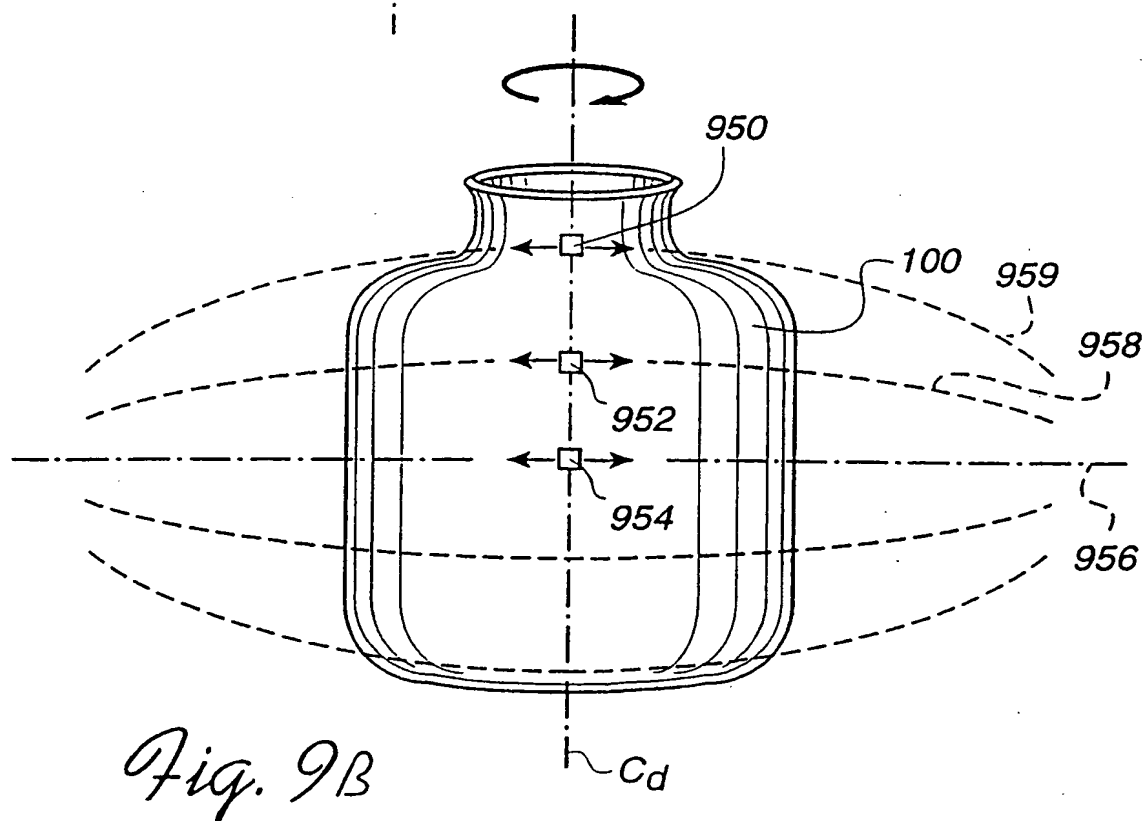
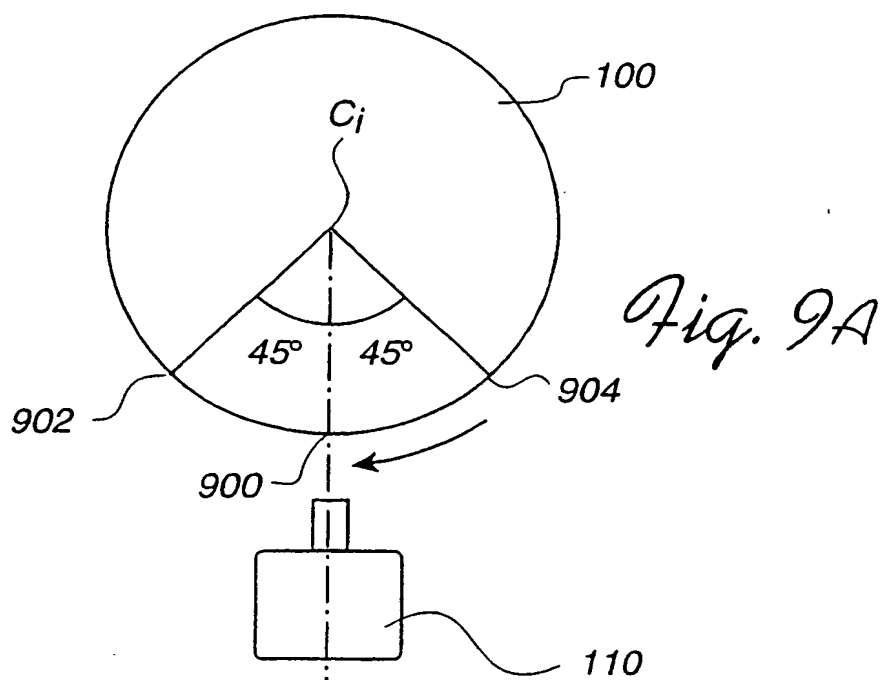
Fig. 5C

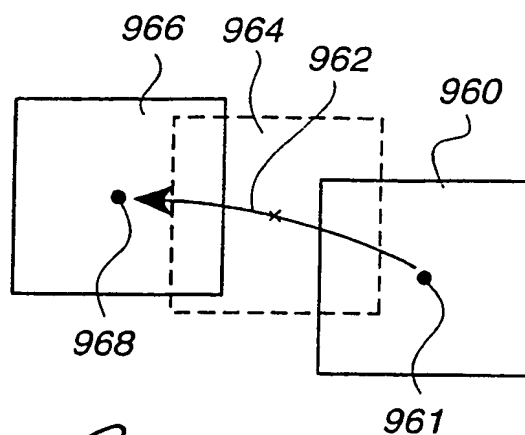


*Fig. 8*

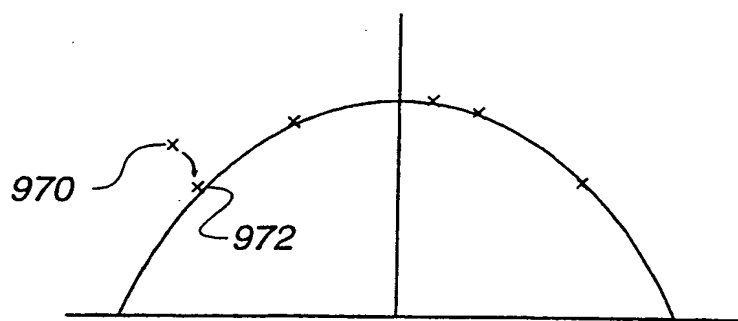
*Fig. 8A*

*Fig. 9*

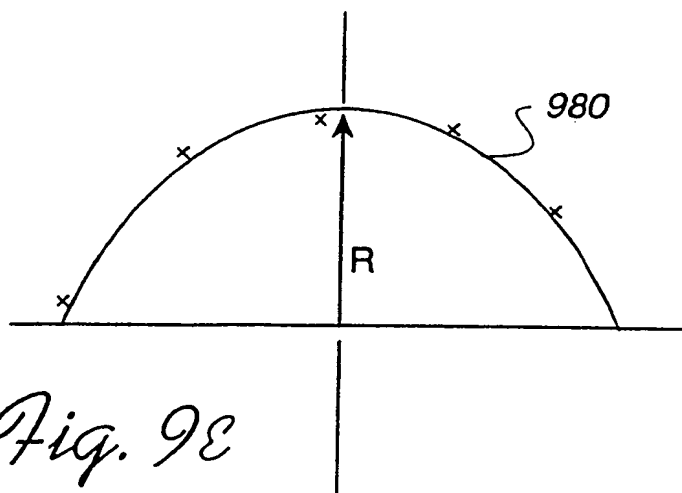




*Fig. 9C*

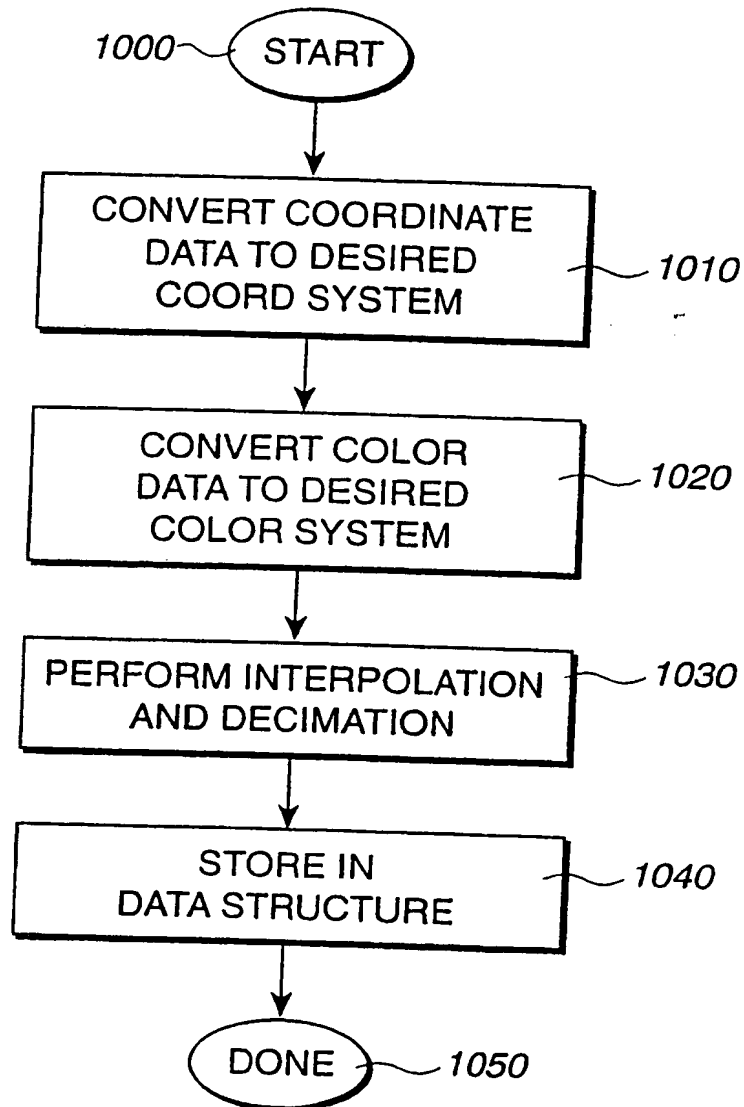


*Fig. 9D*



*Fig. 9E*



*Fig. 10*

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 97/19331

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 G01B11/00 G06T7/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 G01B G06T

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages  | Relevant to claim No.                     |
|------------|---|---|
| X<br>A     | US 4 969 106 A (VOGEL JEFFREY H ET AL) 6<br>November 1990<br><br>see column 2, line 20 - column 3, line 8<br>see column 3, line 58 - column 4, line 41;<br>figure 1<br>see column 5, line 53 - column 6, line 9<br>see column 6, line 56 - column 8, line 34<br>----- | 1-9, 14,<br>16, 17<br>11-13,<br>15, 18-23 |

☐ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

## \* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&amp;" document member of the same patent family

Date of the actual completion of the international search

20 February 1998

Date of mailing of the international search report

02/03/1998

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

De Paepe, W

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 97/19331

| Patent document<br>cited in search report | Publication<br>date | Patent family<br>member(s) | Publication<br>date |
|---|---------------------|----------------------------|---------------------|
| US 4969106 A                              | 06-11-90            | AU 5262990 A               | 26-09-90            |
|   |                     | DE 4090293 T               | 30-01-92            |
|   |                     | JP 4500408 T               | 23-01-92            |
|   |                     | WO 9010194 A               | 07-09-90            |
| <hr/>                                     |                     |                            |                     |



|   |   |  |
|---|---|--|
| (51) 国際特許分類<br>G06T 15/00   | A1  | (11) 国際公開番号<br>WO98/09253<br><br>(43) 国際公開日<br>1998年3月5日(05.03.98) |
| (21) 国際出願番号 PCT/JP97/02997<br>(22) 国際出願日 1997年8月28日(28.08.97)<br>(30) 優先権データ<br>特願平8/248739 1996年8月29日(29.08.96) JP<br>特願平8/248946 1996年8月30日(30.08.96) JP<br>特願平8/248958 1996年8月30日(30.08.96) JP<br>特願平9/214014 1997年7月23日(23.07.97) JP<br>特願平9/217486 1997年8月12日(12.08.97) JP<br>特願平9/217550 1997年8月12日(12.08.97) JP<br>(71) 出願人 (米国を除くすべての指定国について)<br>三洋電機株式会社(SANYO ELECTRIC CO., LTD.)(JP/JP)<br>〒570 大阪府守口市京阪本通2丁目5番5号 Osaka, (JP)<br>(72) 発明者; および<br>(75) 発明者/出願人 (米国についてのみ)<br>松本幸則(MATSUMOTO, Yukinori)(JP/JP)<br>〒305 茨城県つくば市竹園2-16-12-25 Ibaraki, (JP)<br>寺崎 肇(TERASAKI, Hajime)(JP/JP)<br>〒305 茨城県つくば市二の宮4-13-1-2-108 Ibaraki, (JP)<br>杉本和英(SUGIMOTO, Kazuhide)(JP/JP)<br>〒305 茨城県つくば市竹園2-16-12-45 Ibaraki, (JP) | 荒川 勉(ARAKAWA, Tsutomu)(JP/JP)<br>〒301 茨城県竜ヶ崎市長山6-15-19 Ibaraki, (JP)<br>(74) 代理人<br>弁理士 深見久郎, 外(FUKAMI, Hisao et al.)<br>〒530 大阪府大阪市北区南森町2丁目1番29号<br>住友銀行南森町ビル Osaka, (JP)<br>(81) 指定国 CN, US.<br>添付公開書類<br>国際調査報告書 |  |

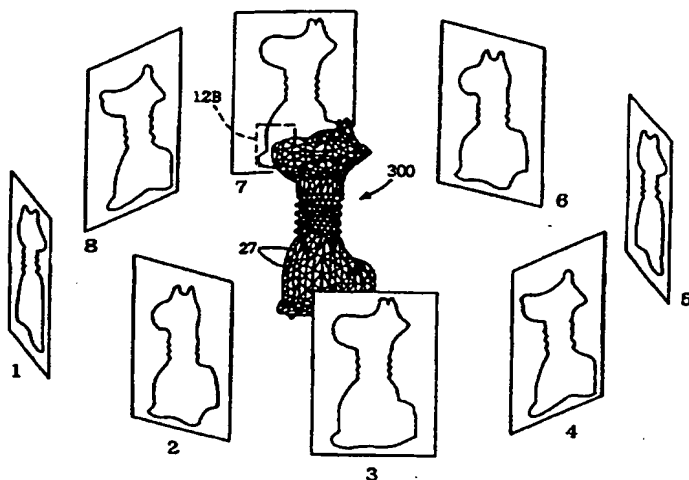
(54) Title: TEXTURE INFORMATION GIVING METHOD, OBJECT EXTRACTING METHOD, THREE-DIMENSIONAL MODEL GENERATING METHOD AND APPARATUS FOR THE SAME

(54) 発明の名称 テクスチャ情報付与方法、物体抽出方法、立体モデル作成方法、およびその装置

(57) Abstract

This method is such that a three-dimensional model of an actual object is expressed by polygons in accordance with information about a plurality of images formed by imaging the actual object while the object is turned in steps of an arbitrary angle, and texture information is given to each polygon from information about the object image whose polygon has the largest projection area. In order to improve the color continuity between adjacent polygons, the object image information corresponding to the polygon concerned and to the polygon adjacent to it are so selected as to be the image information about object images which are picked up from directions close to each other in positions close to each other. Another method is such that an object image is divided into a plurality of regions, the differentials between the object image and the background image is obtained

for each region, the average of the absolute values of the differentials is obtained for each region and regions whose averages of the absolute values of the differentials are larger than a threshold value are detected as object parts. Further, another method is such that only the background of an object is imaged, the object is imaged while it is turned to obtain a plurality of object images, a silhouette image is generated by a differential processing of the object images and the background image, a voting processing of a voxel space is carried out in accordance with the silhouette image, a polygon is generated in accordance with a three-dimensional form obtained by the voting processing, and a texture obtained from the object images is mapped on the polygon.



## (57) 要約

この方法は、現実の物体を任意の角度ごとに回転させて撮影された複数の物体画像情報に基づいて、3次元形状モデルをポリゴンにより表現し、各ポリゴンに対して当該ポリゴンが最も大きな投影面積を有する物体画像情報からテクスチャ情報を付与する。隣接するポリゴン間の色彩連続性を向上させるために、注目するポリゴンと隣接するポリゴンとに対応付けられる物体画像情報は、撮影位置および撮影方向が近接した物体画像情報となるように選択される。もう1つの方法は、物体画像を複数の領域に分割し、領域単位で物体画像と背景画像との差分を求め、領域単位で差分の絶対値の平均値を導出し、差分の絶対値の平均値がしきい値以上の領域を物体部分として検出する。さらにもう1つの方法は、対象物体の背景だけを撮影するとともに、対象物体を回転させながら撮影して複数の物体画像を得る。そして、物体画像と背景画像との間で差分処理を行なうことによってシルエット画像を生成し、そのシルエット画像をもとにボクセル空間にボーディング処理を行ない、ボーディング処理によって得られた立体形状をもとにポリゴンを生成し、物体画像から獲得したテクスチャをポリゴンにマップする。

PCTに基づいて公開される国際出願のパブリック第一頁に記載されたPCT加盟国を特定するために使用されるコード（参考情報）

|    |              |    |             |    |          |    |            |
|----|--------------|----|-------------|----|----------|----|------------|
| AL | アルバニア        | ES | スペイン        | LK | スリランカ    | SE | スウェーデン     |
| AM | アルメニア        | FI | フィンランド      | LR | リベリア     | SG | シンガポール     |
| AT | オーストリア       | FR | フランス        | LS | レソト      | SI | スロベニア共和国   |
| AU | オーストラリア      | GB | 英国          | LT | リトアニア    | SK | スロバキア共和国   |
| AZ | アゼルバイジャン     | GE | グルジア        | LU | ルクセンブルグ  | SL | シエラレオネ     |
| BA | ボスニア・ヘルツェゴビナ | GH | ガーナ         | LV | ラトヴィア    | SN | セネガル       |
| BB | バルバドス        | GM | ガンビア        | MC | モナコ      | SZ | スワジランド     |
| BE | ベルギー         | GN | ギニア         | MD | モルドヴァ共和国 | TD | チャド        |
| BG | ブルガリア        | GW | ギニアビサウ      | MG | マダガスカル   | TG | トーゴ        |
| BJ | ベナン          | GR | ギリシャ        | MK | マケドニア共和国 | TJ | タジキスタン     |
| BR | ブラジル         | HU | ハンガリー       | ML | マリ       | TM | トルクメニスタン   |
| BY | ベラルーシ        | ID | インドネシア      | MN | モンゴル     | TR | トルコ        |
| CA | カナダ          | IE | アイルランド      | MR | モーリタニア   | TT | トリニダード・トバゴ |
| CF | 中央アフリカ共和国    | IL | イスラエル       | MW | マラウイ     | UA | ウクライナ      |
| CG | コンゴ          | IS | アイスランド      | MX | メキシコ     | UG | ウガンダ       |
| CH | スイス          | IT | イタリア        | NE | ニジェール    | US | 米国         |
| CI | コート・ジボアール    | JP | 日本          | NL | オランダ     | UZ | ウズベキスタン    |
| CM | カメルーン        | KE | ケニア         | NO | ノルウェー    | VN | ヴェトナム      |
| CN | 中国           | KG | キルギスタン      | NZ | ニュージーランド | YU | ユーゴスラビア    |
| CU | キューバ         | KP | 朝鮮民主主義人民共和国 | PL | ポーランド    | ZW | ジンバブエ      |
| CZ | チェコ共和国       | KR | 大韓民国        | PT | ポルトガル    |    |            |
| DE | ドイツ          | KZ | カザフスタン      | RO | ルーマニア    |    |            |
| DK | デンマーク        | LC | セント・ルシア     | RU | ロシア連邦    |    |            |
| EE | エストニア        | LI | リヒテンシュタイン   | SD | スーダン     |    |            |

## 明細書

テクスチャ情報付与方法、物体抽出方法、立体モデル作成方法、およびその装置

5

## 技術分野

本発明は、実在する対象物体を撮影して得られた物体画像に基づいてその対象物体の形状モデルにテクスチャ情報を付与するためのテクスチャ情報付与方法、物体画像から背景のような不所望部分を除くようにして物体部分を抽出するための物体抽出方法、対象物体の立体モデルを作成する立体モデル作成方法、およびこれらの装置に関する。

10

## 背景技術

コンピュータグラフィックス等の発達により、3次元グラフィックスの実用的なシステムの開発が盛んに行なわれている。しかしながら、このような実用化システムの普及に伴い顕在化した問題の1つに、形状データの入力方法がある。すなわち、自由曲面を有する物体や、自然界に存在する多くの物体等の複雑な3次元形状を計算機等に入力するのは極めて煩雑で困難な作業となっている。

15

さらに、計算機内等で物体を再構成した場合に、単に物体形状を再構成しただけでは、よりリアルな物体表面の質感等を表現することは困難である。

20

したがって、現実の物体を撮影することにより得た画像情報をもとに、計算機内で形状情報および色彩／質感情報等を再構成することができれば、3次元画像情報の取扱いを簡易に行なうことが可能となる。

25

一方、たとえばインターネット等における3次元画像通信においては、情報発信者たる一般ユーザによって3次元画像を生成する機会が増加する。したがって、簡便かつコンパクトな装置構成で、3次元画像を生成することも必要になる。

(1) 特開平5-135155号公報には、普通の照明条件下で、回転テーブルにある対象物体の連続シルエット画像から3次元モデルを構成することができる立体モデル作成装置が開示されている。

この3次元モデル構成装置は、回転台で回転する対象物体をカメラで連続的に撮像し、画像処理コンピュータによって撮像した画像から対象物体のシルエット画像の抽出を行なう。このシルエット画像において、シルエット画像の輪郭から垂直回転軸までの水平距離を計測し、この水平距離と回転角度とに基づいて3次元モデルが生成される。すなわち、連続的に撮像されたシルエット画像から、対象物体の輪郭を抽出し、3次元モデルとして表示される。

図1は、以上のようにして生成された3次元モデルに対して、カメラにより連続的に撮像された画像情報に基づいて、テクスチャ情報を張付ける方法を示す概念図である。

10 上記特開平5-135155号公報においては、対象物体を連続的に回転して、対象物体を撮像する、言換えると、人間の3次元モデルに対する形状認識の分解能程度で画像情報を得る場合が説明されている。より具体的には、たとえば回転角が $1^\circ$ ごとに画像を撮像し、対象物体に対して、360枚の画像を得る場合が説明されている。

15 以下では、説明を簡単にするために、より粗い角度ごとに、画像を撮像した場合について説明するが、その本質は同等である。

図1に示すように、対象物体を、一定の角度ずつ回転させながら、合計 $n$ 枚の画像を撮像した場合を考える。この場合、各画像情報には、ラベル番号として、 $1, 2, \dots, n$ が対応しているものとする。

20 対象物体は、ポリゴン（三角パッチ）を用いた形状モデル（ワイヤフレームモデル）300で表現されているものとする。この形状モデル300に対してテクスチャ情報を付与する場合、三角パッチの各々に対して、対象物体を撮像したカメラの方向に応じて、対応するラベル番号の画像情報の色彩情報（テクスチャ情報）を付与する。

25 より詳しく説明すると、形状モデル300における回転軸から、注目する三角パッチに向かうベクトルを考えた場合、このベクトルと撮影方向ベクトルとの方向が最も一致するような画像から、この三角パッチに対するテクスチャ情報を獲得するという方法である。あるいは、より直感的に説明すれば、表面モデルに地球儀のような経線を仮定し、 $0^\circ \sim 1 \times 360/n^\circ$ の範囲の三角パッチについ

ては1番の画像情報から、 $1 \times 360 / n^\circ \sim 2 \times 360 / n^\circ$ の範囲の三角パッチについては2番の画像情報から、というようにテクスチャ情報を獲得するという方法である。以下では、このようなテクスチャ情報の獲得方法を中心投射方式と呼ぶことにする。

- 5       以上説明したような中心投射方式では、各三角パッチあるいはこれに相当する形状モデルを構成する構成要素（以下、立体形状構成要素と呼ぶ）に対して、1対1に画像情報を対応させることが可能で、かつ、この対応関係の決定が容易であるという利点がある。

- 10       しかしながら、中心投射方式では、回転軸から見た場合、同一の回転角範囲内に存在しない立体形状構成要素に対しては、異なる画像情報（ラベル番号の異なる画像情報）からテクスチャ情報が付与されるため、照明等の関係により微妙に色彩情報の光沢や質感が異なる場合そのテクスチャに継ぎ目が目立つという欠点がある。

- 15       また、対象物体の形状によっては、ある撮像方向から得た画像情報では、対応する立体形状構成要素が見通せないために、対応する画像情報中に、その立体形状構成要素に対応するテクスチャ情報が全く含まれないという場合が生じる。

- 20       図2は、このような場合を説明するための概念図である。すなわち、図2においては、対象物体の回転軸を含む垂直平面における、回転軸、対象物体断面およびカメラ内に射影される物体像の関係を示す。対象物体の形状が、図2に示すような場合、すなわち、カメラ側から見通すことができない部分を含んでいる場合は、この角度方向から撮像した画像情報には、この見通せない部分のテクスチャ情報が存在しないことになる。ところが、たとえばこの撮像方向と一定の角度を有する別の撮像方向からは、この見通せなかった部分のテクスチャ情報を獲得することが可能である。

- 25       (2) 物体画像から物体部分だけ抽出する従来の方法としては、補助ツールを使った人手によるものがある。具体的には、対象物体を背景とともに撮影して得られた物体画像を複数の領域に分割する。そして、人がマウスなどを用いて、物体画像中の背景部分を選択し、背景部分を消していく。しかしながら、この方法は、人手によるため、多大な労力を要するという問題点がある。



また、従来の別の物体抽出方法としては、クロマキー技術を使ったものがある。具体的には、同一色の背景板を用いて物体画像から物体部分を抽出する。しかしながら、この方法では、同一色の背景板という特殊な環境を用意する必要があるという問題点がある。

- 5       また、従来の別の物体抽出方法としては、単純差分を用いたものがある。具体的には、物体画像と、対象物体の背景だけを撮影した背景画像との間で差分処理を行ない、差分を求める。そして、差分の絶対値がしきい値以上の部分を物体部分として抽出する。しかしながら、この方法では、対象物体に背景の色と同じ部分が存在する場合には、その部分を物体部分として抽出できないという問題点がある。
- 10       つまり、物体部分の抽出精度が悪いという問題点がある。

- また、従来の別の物体抽出方法としては、ステレオ法による奥行情報を利用したものがある。具体的には、対象物体を背景とともに撮影して得られた物体画像から奥行情報がしきい値以下の部分を物体部分として抽出する。しかしながら、この方法では、対象物体と背景との境界付近では奥行の差が大きく、必ずしも正確な奥行情報を得られない。そのため、背景画像の一部分が誤って物体部分として抽出されるという問題点がある。
- 15       また、上記従来の方法はいずれも、しきい値を予め定めておく必要がある。ここで、画像をA/D変換するA/D変換器の変換特性や照明の特性を考慮して、適切なしきい値を定めることは極めて難しいという問題点がある。特に、A/D変換器の変換特性や照明の特性が変動する場合には、改めてしきい値を定め直す必要があるという問題点もある。
- 20

- (3) 対象物体の形状を読取る従来の装置としては、3次元デジタイザがある。3次元デジタイザは、複数の関節を持つアームおよびペンを備える。そして、人が、このペンを対象物体に接触させ、対象物体上を動かしていく。このようにペンを動かすと、アームの関節の角度が変わっていく。このアームの関節の角度情報をもとに、対象物体の3次元形状を求める。しかしながら、このようなデジタイザでは、人手で計測を行なうため、作業時間および労力が非常に大きくなるという問題点がある。
- 25

      また、従来の別の装置としては、レーザスキャナがある。レーザスキャナは、

レーザを対象物体に照射し、スキャンすることにより、対象物体の3次元形状を求めるものである。しかしながら、このようなレーザスキャナでは、光を吸収する材質からなる対象物体の立体モデルを得ることができないという問題点がある。さらに、装置が非常に複雑で高価になるという問題点がある。さらに、暗室で対象物体の計測を行なう必要があり、撮影環境が制限されるという問題点がある。さらに、色情報を取込みにくいという問題点がある。

また、米国特許4,982,438号には立体モデル作成装置が開示されている。この装置では、対象物体のシルエット画像を用いて、仮定存在領域を算出する。仮定存在領域とは、カメラの投影中心を頂点とし、対象物体のシルエットを断面形状とする錐体状の領域である。この錐体状の領域（仮定存在領域）をボクセルモデルで記述する。以上のような処理を複数のシルエット画像に対して行なう。そして、共通仮定存在領域を求め、対象物体の立体モデルを作成する。ここで、共通仮定存在領域とは、複数のシルエット画像に対する複数の仮定存在領域の論理積部分である。しかしながら、この装置では、論理積により立体形状を得ているため、1つのシルエット画像が不正確であった場合、精度よく立体モデルを作成することができないという問題点がある。また、水平方向（回転軸に垂直な方向）からのみ、対象物体を撮影しているため、色情報が不足したり、局所的凹部を認識できない場合があるという問題点がある。

また、上記特開平5-135155号公報に開示された立体モデル作成装置では、回転台で回転する対象物体をカメラで撮影し、複数のシルエット画像を得る。この複数のシルエット画像をもとに、複数の水平面（回転軸に垂直な面）での対象物体の複数の形状を得る。そして、隣り合う水平面において、対象物体の形状の輪郭線上の点同士を三角パッチで結ぶ。ここで、1つの水平面での対象物体の形状の輪郭線上の点は、所定角度ごとに定められる。このようにして、対象物体の立体モデルを作成する。しかしながら、この装置では、シルエット画像を作成するために、背景板を用いており、特殊な撮影環境が必要になるという問題点がある。さらに、複数の水平面での対象物体の形状を用いて立体モデルを作成しているためデータ量が多くなり、処理に時間がかかるという問題点がある。

それゆえに、本発明の目的は、実在する物体を撮像した画像情報に基づいて、

計算機内等において立体モデルを再構成する場合、対象物体の形状にかかわらず、形状モデルを構成する立体形状構成要素の各々にテクスチャ情報を付与することが可能なテクスチャ情報付与方法およびその装置を提供することである。

5 本発明の他の目的は、現実撮像された画像情報に基づいてテクスチャ情報を形状モデルに付与する際に、撮像された画像情報からより現実の物体の有するテクスチャに近いテクスチャ情報を付与することが可能なテクスチャ情報付与方法およびその装置を提供することである。

10 本発明のさらに他の目的は、現実の物体を撮像して得られた画像情報をもとにテクスチャ情報を形状モデルに付与する際に、形状モデルを構成する立体形状構成要素の各々に付与されるテクスチャの不連続性（継ぎ目）が目立ちにくいテクスチャ情報付与方法およびその装置を提供することである。

本発明のさらに他の目的は、対象物体に背景の色と同じ部分が存在する場合でもその部分を物体部分として抽出できる物体抽出方法およびその装置を提供することである。

15 本発明のさらに他の目的は、種々の特性が変動する場合でも、常に安定して正確な物体部分を抽出できる物体抽出方法およびその装置を提供することである。

本発明のさらに他の目的は、人手による作業を少なくするとともに特殊な撮影環境を不要とする物体抽出方法およびその装置を提供することである。

20 本発明のさらに他の目的は、人手による作業を少なくすることのできる立体モデル作成方法およびその装置を提供することである。

本発明のさらに他の目的は、撮影環境および対象物体の材質の制限が小さく、簡単な構成の立体モデル作成方法およびその装置を提供することである。

25 本発明のさらに他の目的は、複数のシルエット画像のうちのいくつかが不正確であっても、精度よく立体モデルを作成できる立体モデル作成方法およびその装置を提供することである。

本発明のさらに他の目的は、十分な色情報を得ることができ、対象物体の局所的凹部をも認識できる立体モデル作成方法およびその装置を提供することである。

本発明のさらに他の目的は、特殊な撮影環境が不要で、処理するデータ量が少なく高速に立体モデルを作成できる立体モデル作成方法およびその装置を提供す

ることである。

#### 発明の開示

本発明の 1 つの局面に従うと、形状モデルに対するテクスチャ情報付与装置は、  
5 対象物体の形状を複数の立体形状構成要素の集合による形状モデルとして記述する手段と、対象物体を異なる視点から撮影することにより獲得された複数の物体画像情報に基づいて、形状モデルに対するテクスチャ情報を、立体形状構成要素ごとに、各物体画像情報の立体形状構成要素に対するテクスチャ情報量に応じて付与する手段とを備える。

10 好ましくは、上記テクスチャ情報量は、立体形状構成要素ごとに、各立体形状構成要素の面法線方向と、各物体画像情報を撮影した方向との一致の程度により表現される。

好ましくは、上記テクスチャ情報量は、立体形状構成要素ごとに、各物体画像情報に投影され立体形状構成要素の面積により表現される。

15 本発明のもう 1 つの局面に従うと、形状モデルに対するテクスチャ情報付与装置は、対象物体の形状を複数の立体形状構成要素の集合による形状モデルとして記述する手段と、対象物体を異なる視点から撮影することにより獲得された複数の物体画像情報に基づいて、形状モデルに対するテクスチャ情報を立体形状構成要素ごとに、各物体画像情報の立体形状構成要素に対するテクスチャ情報量と立  
20 体形状構成要素間のテクスチャ連続性との双方に応じて付与する手段とを備える。

好ましくは、上記テクスチャ情報を付与する手段においては、テクスチャ情報量の増加に伴い減少し、かつ、立体形状構成要素間のテクスチャ連続性の向上に伴い減少する評価関数を極小化するように、立体形状構成要素ごとに対応付けられた物体画像情報から、形状モデルに対するテクスチャ情報を付与する。

25 好ましくは、上記評価関数においては、テクスチャ連続性は、注目する立体形状構成要素とそれに隣接する立体形状構成要素とにそれぞれ対応する物体画像情報の撮影位置および撮影方向の差の関数として表現される。

好ましくは、上記評価関数においては、テクスチャ連続性は、位置変化を伴って物体画像情報を撮影した場合に各物体画像情報に位置変化に対応したラベル番

号を与えたとき、注目する立体形状構成要素に割当てられたラベル番号と、注目する立体形状構成要素に隣接する立体形状構成要素に割当てられたラベル番号との差の増加に伴い増加する関数として表現される。

5 好ましくは、上記評価関数においては、テクスチャ連続性は、規則的な位置変化を伴って物体画像情報を撮影した場合に各物体画像情報に位置変化に対応したラベル番号を与えたとき、注目する立体形状構成要素に割当てられたラベル番号と、注目する立体形状構成要素に隣接する立体形状構成要素に割当てられたラベル番号との差の増加に伴い増加する関数として表現される。

10 好ましくは、上記評価関数においては、テクスチャ情報量は、立体形状構成要素ごとに、各物体画像情報に投影される立体形状構成要素の面積の関数として表現される。

好ましくは、上記評価関数においては、テクスチャ情報量は、立体形状構成要素ごとに、各立体形状構成要素の面法線方向と、各物体画像情報を撮影した方向との一致の程度の関数として表現される。

15 好ましくは、上記評価関数は、 $i$  番目 ( $i$  : 自然数) の立体形状構成要素に割当てられたラベル番号と  $i$  番目の立体形状構成要素に隣接する立体形状構成要素に割当てられたラベル番号との差の全立体形状構成要素についての総和と、 $i$  番目の立体形状構成要素に割当てられたラベル番号に対応する物体画像情報に投影される、 $i$  番目の立体形状構成要素の面積の全立体形状構成要素についての総和との線形結合で表現される。

20

本発明のさらにもう 1 つの局面に従うと、形状モデルに対するテクスチャ情報付与装置は、対象物体の形状を複数の立体形状構成要素の集合による形状モデルとして記述する手段と、位置変化を伴って複数の物体画像情報を撮影した場合に各物体画像情報に位置変化に対応したラベル番号を与えたとき、各立体形状構成要素に対するテクスチャ情報量の増加に伴い減少し、かつ、各立体形状構成要素とそれに隣接する立体形状構成要素とにそれぞれ割当てられるテクスチャ情報のテクスチャ連続性の向上に伴い減少する評価関数を極小化するように、立体形状構成要素ごとにラベル番号を対応付ける手段と、対応付けられたラベル番号に対応する物体画像情報と、対応付けられたラベル番号を含む所定数のラベル番号に

25

対応する物体画像情報とに基づいて、各物体画像情報に投影される立体形状構成要素の面積に応じた重み平均処理を行なうことにより、立体形状構成要素にテクスチャ情報を付与する手段とを備える。

- 好ましくは、上記立体形状構成要素にテクスチャ情報を付与する手段では、上記立体形状構成要素について、立体形状構成要素に対応づけられたラベル番号に対応する物体画像情報、および上記対応づけられたラベル番号を含む所定数のラベル番号に対応する物体画像情報に投影された面積を求め、これを下記重み平均処理を行なう際の重み係数とする。そして、上記立体形状構成要素のテクスチャ情報として、上記立体形状構成要素が上記物体画像情報に投影される部位を求め、この投影部の画像情報（色、濃度あるいは輝度）を重み平均処理することで上記テクスチャ情報とする。

- 本発明のさらにもう1つの局面に従うと、形状モデルに対するテクスチャ情報付与装置は、対象物体の形状を複数の立体形状構成要素の集合による形状モデルとして記述する手段と、規則的な位置変化を伴って複数の物体画像情報を撮影した場合に各物体画像情報に位置変化に対応したラベル番号を与えたとき、各立体形状構成要素に対するテクスチャ情報量の増加に伴い減少し、かつ、各立体形状構成要素とそれに隣接する立体形状構成要素とにそれぞれ割当てられるテクスチャ情報のテクスチャ連続性の向上に伴い減少する評価関数を極小化するように、立体形状構成要素ごとにラベル番号を対応付ける手段と、対応付けられたラベル番号に対応する物体画像情報と、対応付けられたラベル番号を含む所定数のラベル番号に対応する物体画像情報とに基づいて、各物体画像情報に投影される立体形状構成要素の面積に応じた重み平均処理を行なうことにより、立体形状構成要素にテクスチャ情報を付与する手段とを備える。

- 好ましくは、上記立体形状構成要素にテクスチャ情報を付与する手段では、上記立体形状構成要素について、立体形状構成要素に対応づけられたラベル番号に対応する物体画像情報、および上記対応づけられたラベル番号を含む所定数のラベル番号に対応する物体画像情報に投影された面積を求め、これを下記重み平均処理を行なう際の重み係数とする。そして、上記立体形状構成要素のテクスチャ情報として、上記立体形状構成要素が上記物体画像情報に投影される部位を求め、

この投影部の画像情報（色、濃度あるいは輝度）を重み平均処理することで上記テクスチャ情報とする。

5 本発明のさらにもう1つの局面に従うと、形状モデルに対するテクスチャ情報付与装置は、対象物体を異なる視点から撮影することにより複数の物体画像情報を獲得する手段と、対象物体の形状を複数の立体形状構成要素の集合による形状モデルとして記述する手段と、立体形状構成要素ごとに、複数の物体画像情報のそれぞれに投影された立体形状構成要素に対応する面積に応じた、すべての物体画像情報についての重み平均処理を行なうことにより得られたテクスチャ情報を付与する手段とを備える。

10 好ましくは、上記立体形状構成要素にテクスチャ情報を付与する手段では、上記立体形状構成要素の各々について、物体画像情報に投影された面積を求め、これを下記重み平均処理を行なう際の重み係数とする。そして、上記立体形状構成要素のテクスチャ情報として、上記立体形状構成要素が上記物体画像情報に投影される部位を求め、この投影部の画像情報（色、濃度あるいは輝度）を重み平均  
15 処理することで上記テクスチャ情報とする。

したがって、上記テクスチャ情報付与装置によれば、実在する物体を撮影した画像情報に基づいて、計算機内等において形状モデルを再構成した場合に、当該対象物体を撮影した複数の画像情報のうちから、最も現実の物体が有するテクスチャ情報として最も適切なものをを選択的に形状モデルに付与することが可能で  
20 ある。

また、複数の立体形状構成要素の集合として表現された形状モデルに対して、テクスチャ情報（色彩情報）を付与する場合に、各立体形状構成要素間でのテクスチャ情報の不連続性を抑制しつつ、同時に現実の物体の有するテクスチャ情報により近いテクスチャ情報を各立体形状構成要素に選択的に付与することが可能  
25 である。

さらに、複数の立体形状構成要素に対して、現実の対象物体から撮影した物体画像情報に基づいて、テクスチャ情報を付与する処理を、各立体形状構成要素に対するラベリング問題に置換して処理することが可能であるため、コンピュータ処理等に適した手順で各立体形状構成要素へのテクスチャ情報の付与処理を行な

うことが可能である。

本発明のさらにもう 1 つの局面に従うと、対象物体を撮影して得られた物体画像から不所望部分を除くようにして物体部分を抽出する物体抽出装置は、領域分割手段と、抽出手段とを備える。領域分割手段は、物体画像を複数の領域に分割する。抽出手段は、物体画像の各画素の情報を、領域ごとにまとめる処理を行なうことにより、物体画像中の物体部分を特定し、抽出する。ここで、不所望部分とはたとえば背景部分である。

好ましくは、抽出手段において、物体画像の各画素の情報を、領域ごとにまとめる処理は、物体画像の各画素の情報を、領域ごとに平均化する処理である。

10 好ましくは、抽出手段は、領域ごとにまとめられた各画素の情報に対し、しきい値処理を行なうことにより、物体画像中の物体部分を特定し、抽出する。

好ましくは、物体画像の各画素の情報は、対象物体の背景だけを撮影した背景画像と、物体画像との間で差分処理を行なって得られた差分情報である。

15 好ましくは、抽出手段は、差分処理手段と、平均値導出手段と、しきい値処理手段とを含む。差分処理手段は、対象物体の背景だけを撮影した背景画像と、物体画像との間で差分処理を行なう。平均値導出手段は、差分処理によって得られた差分の絶対値に関して、各領域内の平均値を求める。しきい値処理手段は、領域内の平均値と、予め定められた値とを比較し、平均値が予め定められた値以上の領域を物体部分として抽出する。

20 好ましくは、抽出手段は、平均値導出手段と、差分処理手段と、しきい値処理手段とを含む。平均値導出手段は、物体画像の各領域内の画素の平均値を算出する。差分処理手段は、物体画像の各領域内の画素の平均値とその領域に対応する背景画像の領域内の画素の平均値との間で差分処理を行なう。しきい値処理手段は、差分処理によって得られた差分の絶対値と、予め定められた値とを比較し、  
25 差分の絶対値が予め定められた値以上の領域を物体部分として抽出する。

好ましくは、物体画像の各画素の情報は奥行情報である。

本発明のさらにもう 1 つの局面に従うと、対象物体を撮影して得られた物体画像から、不所望部分を除くようにして物体部分を抽出する物体抽出装置は、奥行情報算出手段と、領域分割手段と、平均値算出手段と、抽出手段とを備える。奥



行情報算出手段は、物体画像の奥行情報を算出する。領域分割手段は、物体画像を複数の領域に分割する。平均値算出手段は、奥行情報の平均値を各領域ごとに算出する。抽出手段は、複数の領域のうち平均値が予め定められた範囲内の領域、特に対象物体よりも前方に位置する物が物体画像中に含まれていない場合は、その平均値が予め定められた値よりも小さい領域を物体部分として抽出する。

本発明のさらにもう１つの局面に従うと、対象物体を撮影して得られた物体画像と、対象物体の背景だけを複数回撮影して得られた複数の背景画像とに基づいて、物体画像から不所望部分を除くようにして物体部分を抽出する物体抽出装置は、差分手段と、抽出手段と、しきい値決定手段とを備える。差分手段は、物体画像と背景画像との差を算出する。抽出手段は、物体画像のうち差がしきい値よりも大きい部分を物体部分として抽出する。しきい値決定手段は、複数の背景画像の分布に基づいてしきい値を統計的に決定する。

本発明のさらにもう１つの局面に従うと、対象物体を撮影して得られた物体画像と、対象物体の背景だけを複数回撮影して得られた複数の背景画像とに基づいて、物体画像から不所望部分を除くようにして物体部分を抽出する物体抽出装置であって、算出手段と、差分手段と、抽出手段とを備える。算出手段は、複数の背景画像において互いに同じ座標に位置する画素の平均値および標準偏差を各画素ごとに算出する。差分手段は、物体画像の各画素の値とその画素に対応する背景画像の画素の平均値との差を算出する。抽出手段は、物体画像の画素のうち差が標準偏差の所定倍よりも大きい画素を物体部分として抽出する。

本発明のさらにもう１つの局面に従うと、対象物体を撮影して得られた物体画像と、対象物体の背景だけを複数回撮影して得られた複数の背景画像とに基づいて、物体画像から不所望部分を除くようにして物体部分を抽出する物体抽出装置は、平均／標準偏差算出手段と、領域分割手段と、差分手段と、平均差算出手段と、平均標準偏差算出手段と、抽出手段とを備える。平均／標準偏差算出手段は、複数の背景画像において互いに同じ座標に位置する画素の平均値および標準偏差を各画素ごとに算出する。領域分割手段は、物体画像を複数の領域に分割する。差分手段は、物体画像の各領域内の各画素の値とその領域に対応する背景画像の領域内の対応する画素の平均値との差を算出する。平均差算出手段は、差の平均

値を各領域ごとに算出する。平均標準偏差算出手段は、標準偏差の平均値を各領域ごとに算出する。抽出手段は、複数の領域のうち差の平均値が標準偏差の平均値の所定倍よりも大きい領域を物体部分として抽出する。

5 本発明のさらにもう1つの局面に従うと、対象物体を撮影して得られた物体画像と、対象物体の背景だけを複数回撮影して得られた複数の背景画像とに基づいて、物体画像から不所望部分を除くようにして物体部分を抽出する物体抽出装置は、平均／標準偏差算出手段と、領域分割手段と、平均算出手段と、差分手段と、平均差算出手段と、平均標準偏差算出手段と、抽出手段とを備える。平均／標準偏差算出手段は、複数の背景画像において互いに同じ座標に位置する画素の平均  
10 値および標準偏差を各画素ごとに算出する。領域分割手段は、物体画像を複数の領域に分割する。平均算出手段は、物体画像の各領域内の画素の平均値を算出する。差分手段は、物体画像の各領域内の画素の平均値とその領域に対応する背景画像の領域内の画素の平均値との差の絶対値を算出する。平均差算出手段は、差の絶対値の平均値を各領域ごとに算出する。平均標準偏差算出手段は、標準偏差  
15 の平均値を各領域ごとに算出する。抽出手段は、複数の領域のうち差の絶対値の平均値が標準偏差の平均値の所定倍よりも大きい領域を物体部分として抽出する。

本発明のさらにもう1つの局面に従うと、対象物体を撮影して得られた物体画像と、対象物体の背景だけを複数回撮影して得られた複数の背景画像とに基づいて、物体画像から不所望部分を除くようにして物体部分を抽出する物体抽出装置  
20 は、平均／標準偏差算出手段と、領域分割手段と、平均算出手段と、差分手段と、平均標準偏差算出手段と、抽出手段とを備える。平均／標準偏差算出手段は、複数の背景画像において互いに同じ座標に位置する画素の平均値および標準偏差を各画素ごとに算出する。領域分割手段は、物体画像を複数の領域に分割する。平均算出手段は、物体画像の各領域内の画素の平均値を算出するとともに、背景画像の画素の平均値の各領域内における平均値を算出する。差分手段は、物体画像  
25 の各領域内の画素の平均値とその領域に対応する背景画像の領域内の画素の平均値の領域内における平均値との差の絶対値を算出する。平均標準偏差算出手段は、標準偏差の平均値を各領域ごとに算出する。抽出手段は、複数の領域のうち差の絶対値が標準偏差の平均値の所定倍よりも大きい領域を物体部分として抽出する。

本発明のさらにもう 1 つの局面に従うと、対象物体を複数回撮影して得られた複数の物体画像と、対象物体の背景だけを複数回撮影して得られた複数の背景画像とに基づいて、物体画像から不所望部分を除くようにして物体部分を抽出する物体抽出装置は、平均／標準偏差算出手段と、平均算出手段と、領域分割手段と、  
5 差分手段と、平均差算出手段と、平均標準偏差算出手段と、抽出手段とを備える。平均／標準偏差算出手段は、複数の背景画像において互いに同じ座標に位置する画素の平均値および標準偏差を各画素ごとに算出する。平均算出手段は、複数の物体画像において互いに同じ座標に位置する画素の平均値を各画素ごとに算出する。領域分割手段は、物体画像を複数の領域に分割する。差分手段は、物体画像  
10 の各領域内の各画素の平均値とその領域に対応する背景画像の領域内の対応する画素の平均値との差の絶対値を算出する。平均差算出手段は、差の絶対値の平均値を各領域ごとに算出する。平均標準偏差算出手段は、標準偏差の平均値を各領域ごとに算出する。抽出手段は、複数の領域のうち差の絶対値の平均値が標準偏差の平均値の所定倍よりも大きい領域を物体部分として抽出する。

したがって、上記物体抽出装置によれば、対象物体に背景の色と同じ部分が存在する場合でもその部分を物体部分として検出し抽出できる。また、人手による作業を少なくできるとともに特殊な撮影環境が不要である。

本発明のさらにもう 1 つの局面に従うと、対象物体の立体モデルを作成する立体モデル作成装置は、対象物体の背景を撮影し、かつ、背景を含めて対象物体を  
20 撮影する撮影手段と、背景だけを撮影して得られた背景画像と、背景を含めて対象物体を撮影して得られた複数の物体画像との差分を求め、複数のシルエット画像を作成するシルエット作成手段と、複数のシルエット画像を用いて、対象物体の立体モデルを作成する手段とを備える。

好ましくは、上記立体モデル作成装置は、対象物体を回転させるための回転手段をさらに備える。  
25

本発明のさらにもう 1 つの局面に従うと、対象物体の立体モデルを作成する立体モデル作成装置は、対象物体の複数のシルエット画像を作成するシルエット作成手段と、複数のシルエット画像をもとにボクセル空間における対象物体の存在領域を推定する推定手段と、推定手段によって得られた対象物体の存在領域を用

いて、対象物体の立体モデルを作成する手段とを備える。

好ましくは、上記推定手段はボクセル空間へボーディング処理を行なう。

好ましくは、上記立体モデル作成装置は、ボーディング処理の結果、投票数が所定のしきい値以上の部分を、対象物体の存在領域とするしきい値処理手段をさらに備える。

したがって、上記立体モデル作成装置によれば、差分処理を行なうことによって得られたシルエット画像を用いて立体モデルを作成するため、同一色の背景板などのような特殊な撮影環境が不要である。

また、複数のシルエット画像をもとにボクセル空間へボーディング処理を行なうことによって立体モデルを作成するため、複数のシルエット画像のうちのいくつかが不正確であっても精度よく立体モデルを作成できる。

また、対象物体の立体形状を切断して得られた複数の切断面の輪郭線を多角形近似することによって立体モデルを作成するため、立体モデル作成のためのデータ量を少なくすることができ、高速処理が可能になる。

また、対象物体の複数の断面形状の輪郭線を多角形近似することによって立体モデルを作成するため、立体モデル作成のためのデータ量を少なくすることができ、高速処理が可能になる。

#### 図面の簡単な説明

図 1 は、従来のテクスチャ情報付与方法を示す概念図である。

図 2 は、従来のテクスチャ情報付与方法の問題点を示す断面図である。

図 3 は、本発明の実施の形態 1 の立体モデル作成装置の構成を示す概略ブロック図である。

図 4 は、図 3 に示された立体モデル作成装置における色彩情報付与処理部の構成を示す概略ブロック図である。

図 5 は、現実の物体から立体モデルを作成する処理のフローを示すフローチャートである。

図 6 A は、図 5 のステップ S 1 0 における画像撮影を説明するための図である。

図 6 B は、図 5 のステップ S 1 2 におけるシルエット画像生成を説明するため

の図である。図 6 C は、図 5 のステップ S 1 4 におけるボーディング処理を説明するための図である。図 6 D は、図 5 のステップ S 1 6 におけるポリゴン生成を説明するための図である。図 6 E は、図 5 のステップ S 1 8 におけるテクスチャマッピングを説明するための図である。

5 図 7 は、ボーディング処理の概念を示す透視図である。

図 8 は、ボーディング処理における仮定存在領域を示す P 平面断面図である。

図 9 はボーディング処理の概念を示す P 平面断面図である。

図 1 0 A は、ポリゴン生成処理を説明するための概念図である。図 1 0 B は、図 1 0 A 中の 1 0 B 部の拡大図である。

10 図 1 1 は、ポリゴンによる 3 次元形状モデルを示す図である。

図 1 2 A は、テクスチャ情報の付与処理を示す概念図である。図 1 2 B は、図 1 2 A 中の 1 2 B 部の拡大図である。

図 1 3 は、各立体形状構成要素へのテクスチャ情報付与の処理の流れを示すフローチャートである。

15 図 1 4 は、実施の形態 1 のテクスチャ情報付与方法の処理の流れを示すフローチャートである。

図 1 5 は、実施の形態 1 のテクスチャ情報付与方法を記録した記録媒体を示す概念図である。

20 図 1 6 は、実施の形態 2 のテクスチャ情報付与方法の処理の流れを示すフローチャートである。

図 1 7 は、実施の形態 3 のテクスチャ情報付与方法の処理の流れを示すフローチャートである。

図 1 8 は、図 4 中の色彩情報記憶部へのテクスチャ情報の記憶方法を示す概念図である。

25 図 1 9 は、実施の形態 4 のテクスチャ情報付与方法の処理の流れを示すフローチャートである。

図 2 0 は、実施の形態 5 のテクスチャ情報付与方法を示す概念図である。

図 2 1 は、実施の形態 5 のテクスチャ情報付与方法の処理の流れを示すフローチャートである。

図 2 2 は、本発明の実施の形態 6 による物体抽出装置（画像切出装置）の全体構成を示す図である。

図 2 3 は、本発明の実施の形態 6 による物体抽出装置（画像切出装置）の概略を示すブロック図である。

5 図 2 4 は、図 2 2 の演算部の概略を示すブロック図である。

図 2 5 A から図 2 5 C は、図 2 4 の差分処理部、平均値導出部およびしきい値処理部における処理を詳細に説明するための図である。

図 2 6 は、本発明の実施の形態 7 による物体抽出装置の主要構成を示すフローチャートである。

10 図 2 7 A は、図 2 6 に示された物体抽出装置において得られる複数の領域 R に分割された物体画像を示す図である。図 2 7 B は、奥行情報を輝度で表示する画像を示す図である。図 2 7 C は物体画像から背景部分を除いて抽出された物体部分の画像を示す図である。

15 図 2 8 は、実施の形態 8 による物体抽出装置の主要構成を示すフローチャートである。

図 2 9 は、実施の形態 9 による物体抽出装置の主要構成を示すフローチャートである。

図 3 0 は、実施の形態 1 0 による物体抽出装置の主要構成を示すフローチャートである。

20 図 3 1 は、実施の形態 1 1 による物体抽出装置の主要構成を示すフローチャートである。

図 3 2 は、実施の形態 1 2 による立体モデル作成装置の概略を示すブロック図である。

25 図 3 3 は、図 3 2 に示された立体モデル作成装置における処理の流れを示すフローチャートである。

図 3 4 は、図 3 3 のステップ S 8 で求められるパースペクティブ比を説明するための図である。

図 3 5 A から図 3 5 C は、図 3 3 のステップ S 8 で求められるカメラと回転テーブルの位置関係を説明するための図である。

図36は、図33のステップS14で用いられる円筒座標系ボクセル空間のボクセルを説明するための図である。

図37は、図33のステップS14におけるボーテイング処理を説明するための図である。

5 図38は、図33のステップS14におけるボーテイング処理の結果を示す図である。

図39Aは、図33のステップS16におけるポリゴン作成の具体的内容を説明するための図である。図39Bは、図39A中の39B部の拡大図である。

10 図40は、図33のステップS16におけるポリゴン作成の流れを説明するための図である。

図41は、図40のステップSA2で求めた、隣接する切断面の輪郭線に対応する頂点間の関係を示す図である。

図42は、図40のステップSA3における局所的最近接点接続戦略を説明するための図である。

15 図43は、図40のステップSA3における局所的最近接点接続戦略によって得られたポリゴンを示す図である。

図44は、図40のステップSA3における局所的最近接点接続戦略によるポリゴン生成の流れの一部を示す図である。

20 図45は、図40のステップSA3における局所的最近接点接続戦略によるポリゴン生成の流れの他の一部を示す図である。

図46は、図40のステップSA3における大域的最短接続戦略によるポリゴン生成の流れを説明するための図である。

図47は、図3のコンピュータに対象物体の立体モデルを作成させるためのプログラムを記録したCD-ROMを示す図である。

25

発明を実施するための最良の形態

以下、本発明の実施の形態を図面を参照して詳しく説明する。なお、図中同一または相当部分には同一符号を付してその説明は繰り返さない。

[実施の形態1]

図3は、本発明の実施の形態1の実在の物体から3次元モデルを再構成するための立体モデル作成装置1000の構成を示す概略ブロック図である。図3を参照して、回転台110の上に、対象物体100が載せられる。回転台110は、たとえば、コンピュータ130からの制御信号に応じて、その回転角度が制御される。カメラ120は、回転する対象物体100を各指定角度ごとに撮影し、得られた画像データをコンピュータ130に与える。一方で、コンピュータ130には、回転台110の回転ピッチ等の撮影条件のデータが入力装置140から与えられる。

コンピュータ130は、カメラ120から与えられた画像情報に基づいて、各撮影角度に対応する画像情報からシルエット画像を抽出し、3次元形状モデルを生成する。ここで3次元形状モデルとしては、たとえばポリゴン（三角パッチ）の集合で表現することが可能である。また、上述した画像情報とは、たとえば、カメラ120から出力される各画素に対応した輝度、色彩、あるいは濃淡を表す数値情報を意味する。ただし、立体モデルの表現方法としては、このような表現方法に限定されることなく、たとえば、異なる形状の表面形状要素の集合として表現することも可能である。したがって、以下では形状モデルを表現するための要素となる形状を立体形状構成要素と総称することにする。

続いて、コンピュータ130は、各角度において撮影された画像情報に基づいて、再構成された形状モデルに対してテクスチャ情報を付与する。ここで、CG（コンピュータグラフィックス）の分野における色彩情報（テクスチャ情報）とは、物体表面に貼り付けられて、この表面の細かな凸凹や絵柄、模様、材質感を表わすための画像情報のことである。このようにして再構成された立体モデルは、表示装置150に表示される。

3次元画像情報に対する色彩情報（より一般的には、テクスチャ情報）の付与方法について詳しく説明する前に、まず実在の物体から3次元形状モデルを生成するまでのフローについて簡単に説明する。

図5は、物体画像を作成してから、形状モデルにテクスチャ情報を付与するまでの処理のフローを示すフローチャートである。図6Aから図6Eは、各フローの過程におけるデータ処理の概要を示す概念図である。



図 6 A を参照して、まず、コンピュータ 130 は、入力装置 140 から与えられた撮影条件データに基づいて回転テーブル 110 の回転角を制御する。各回転角ごとにカメラ 120 が撮像した物体画像 A1 ~ An の取込みが行なわれる（ステップ S10）。たとえば、角度が 10° ごとにカメラ 120 からの物体画像の取込が行なわれるとすると、1 回転で 36 枚の物体画像 A1 ~ An が取込まれる。

続いて図 6 B を参照して、コンピュータ 130 は、撮影された物体画像 A1 ~ An の各々から物体像の輪郭を抽出し、各方向から見た物体のシルエット画像 B1 ~ Bn を生成する（ステップ S12）。

ここで得られたシルエット画像 B1 ~ Bn は、各方向から見た場合の対象物体 100 の輪郭を示す。したがって、図 6 C に示されるように、これら各方向から見た物体の輪郭像に基づいて、仮想的なボクセル（voxel）に分割された 3 次元空間に対して、後に説明するようなボータィング処理を行なうことで、このボクセル空間 251 内で対象物体 100 の存在領域が推定される（ステップ S14）。

続いて図 6 D を参照して、このボクセル空間 251 で表現された物体領域を、ポリゴン（三角パッチ）27 を用いた形状モデル 300 で表現するように変換が行なわれる（ステップ S16）。ここでは、たとえば表現に必要なポリゴン数を少なく抑え、かつ、表現形状の精度を維持することが要求される。このためには、たとえば以下のような方法でポリゴン 27 の生成を行なうことが可能である。

すなわち、まず、円筒座標系で表されたボクセル空間 251 において、円筒座標系の  $\theta$  表面での切断面輪郭線を多角形近似し、ポリゴン 27 の頂点を決定する。続いて、各頂点をそれぞれ最近接の 3 頂点と接続することで三角パッチを生成する。

続いて図 6 E を参照して、生成されたポリゴン 27 の各三角パッチに対して、ステップ S10 において撮影された画像情報からテクスチャ情報を付与する（ステップ S18）。

以上のようなフローにより現実の物体をカメラ 120 で撮影した画像情報をもとに、コンピュータ 130 内で立体モデル 27 が再構成されることになる。

なお、以上の説明では、回転テーブル 110 上に対象物体 100 を載せ、カメ

ラ 1 2 0 は固定した状態で回転テーブル 1 1 0 を回転させ画像を撮影する構成としたが、画像情報を獲得する方法はこのような構成に限定されない。

たとえば、対象物体 1 0 0 は静止したままで、カメラ 1 2 0 を人間が持って移動し、複数の視点から対象物体 1 0 0 を撮影することで、画像データを得る構成とすることも可能である。この場合でも、各画像の撮影を行なった際のカメラ 1 2 0 の位置および撮影を行なう方向がわかれば、これらの情報をもとに以下に述べるのと同様の方法により 3 次元形状モデル 3 0 0 の再構成およびテクスチャ情報の付与を行なうことが可能である。

以下では、さらに、図 5 に示した各処理ステップについて詳しく説明する。

#### 10 (画像撮影およびシルエット画像生成)

上述のとおり、図 3 に示したような構成で、画像撮影を行なう場合は、対象物体 1 0 0 を回転テーブル 1 1 0 に載せ、テーブル 1 1 0 を回転しつつ物体画像 A 1 ~ A n を複数枚撮像する。これに加えて、次のステップ S 1 2 のシルエット画像抽出のため、背景画像の撮影も行なう。

15 続いて、対象物体画像 A 1 ~ A n と背景画像間の差分処理により、対象物体のみを切出したシルエット画像 B 1 ~ B n を生成する。

このとき、背景画像を単一色とするような特殊な撮影環境を不要とし、かつ、安定したシルエット画像生成を可能とするために、単純な画像間差分ではなく、以下に説明するような領域分割処理を加えた画像差分を行なうことも可能である。

20 すなわち、物体画像 A 1 ~ A n を領域分割し、続いて領域単位で背景画像との差分処理を行なう。ここで、差分処理とは、撮像された物体画像情報を、画素ごとにその信号強度について差分を算出する処理を意味する。さらに、領域単位での差分平均をしきい値処理し物体部分の抽出を行なう。

25 以上説明したような方法では、画素レベルで背景と同色のものが存在しても、領域レベルにおいて背景と異なる色があれば、一般には、物体部分として検出できるという性質を持っている。したがって、生成シルエット画像精度の向上を図ることが可能となる。

#### (ボーディング処理)

以上のようにして、対象物体 1 0 0 を、複数の視点から撮影したシルエット画

像情報を得ることができる。これら複数のシルエット画像情報から、物体の3次元形状を再構成するために、以下に説明するようなボーディング処理が行なわれる。

以下では、まず3次元形状を記述するためのボクセルモデルおよびボクセル空間251について図7を参照して説明をする。

ボクセルモデルとは3次元形状を立体的な格子点の有無で記述するモデルであり、ボクセルが定義される空間をボクセル空間251と呼ぶ。ボクセル空間251は認識物体を包含するような、大きさ・位置で配置される。以下では、対象物体100を回転させながら画像撮影を行なったことに対応して、より自然に対象物体形状を表現することが可能な円筒座標系でこのボクセル空間251を表現することにする。

したがって、各々のボクセルは、円筒座標の半径方向の座標を $r$ 、角度方向の座標を $\theta$ 、軸方向の座標を $z$ とすると、 $r$ 、 $\theta$ 、 $z$ を等間隔で分割した場合の体積要素を意味することになる。この体積要素の集合で3次元形状を表現するものがボクセルモデルである。

次に、シルエット画像 $B_1 \sim B_n$ からこのボクセルモデルを再構成するための手順について簡単に説明する。

まず、1つのシルエット画像に基づいて、ボクセル空間251内に、対象物体に対する仮定存在領域50の算出を行なう。ここで、仮定存在領域50とは、図7に示すようにカメラ120の投影中心51を頂点とし、画像の物体像を断面形状とする錐体状の領域を意味する。すなわち、対象物体100は、この領域の内側に必ず存在していることになる。

ボーディング処理とは、この仮定存在領域50内に存在する各々のボクセルに対して、たとえば1つのシルエット画像に対する仮定存在領域50に該当する場合に、数値1を割振る（投票する）という処理を意味する。

図8は、図7に示した $z$ 軸に垂直な平面 $P$ 上のシルエット画像および円筒ボクセル空間251の断面を示す図である。

上述のとおり、カメラ120の投影中心51とする錐体状の領域が仮定存在領域50に該当するため、この領域50に存在する円筒ボクセル空間251内の各

ボクセルに対して数値 1 が割振られる。

図 9 は、複数のシルエット画像 B 1 ～ B n に基づいてボーディング処理を行なった場合について、平面 P 断面の円筒ボクセル空間 2 5 1 を示す断面図である。

図 9 においては、5 つの視点から撮影されたシルエット画像 B 1 ～ B 5 に基づいて円筒ボクセル空間 2 5 1 に対してボーディング処理を行なった場合を示している。この場合、各々のシルエット画像に基づいたボーディング処理は、それぞれのシルエット画像に対する仮定存在領域 5 0 に対して数値 1 を割振ることになるため、5 つのシルエット画像 B 1 ～ B 5 に基づいて、ボーディング処理を行なった場合、図 9 中のクロスハッチで示した領域は、すべてのシルエット画像 B 1 ～ B 5 に基づく仮定存在領域 5 0 が重なっていることになる。言換えると、このクロスハッチされた領域内のボクセルには、それぞれ 5 つのシルエット画像 B 1 ～ B 5 に基づくボーディング処理の結果、数値 5 が割振られていることになる。

したがって、たとえば円筒ボクセル空間 2 5 1 内のボクセルのうち、5 以上が割当てられているボクセルのみを抽出すれば、対象物体 1 0 0 がこの円筒ボクセル空間 2 5 1 内で存在する領域を算出することが可能となる。

より一般的には、撮像された物体画像数に応じて、適当なしきい値を設定することで、ボーディング処理に基づいて円筒ボクセル空間 2 5 1 内における対象物体の存在領域を算出することが可能となる。以上の処理により、円筒ボクセル空間 2 5 1 内において対象物体 1 0 0 の存在する領域を抽出することが可能となる。

ボクセル空間 2 5 1 を用いた 3 次元モデル作成方法としては、他に米国特許 4, 9 8 2, 4 3 8 号に開示された錐体相関法があるが、この方法は、生成するシルエット画像の誤差がそのまま再構成される物体形状に影響を与えてしまうという問題がある。これに対して、ボーディング処理による 3 次元モデル作成方法は、基礎となるシルエット画像に誤りが含まれている場合でも、適当なしきい値を設定することで、獲得される 3 次元形状の精度低下を小さく抑えることが可能であるという特徴を有する。

#### (ポリゴン生成)

続いて、ボクセル空間 2 5 1 で表現された物体領域は、ポリゴン（三角パッチ）2 7 を用いた形状モデル 3 0 0 で表現するように変換される。

図 10 A および図 10 B は、このようなポリゴン生成の過程を示す概念図である。図 10 A および図 10 B を参照して、円筒ボクセル空間 251 で表現された物体領域の円筒座標系  $\theta_1$  平面（円筒座標系で  $\theta = \theta_1$  である平面）での切断面輪郭線を多角形近似する。この多角形近似により得られた輪郭線 L01 の各頂点が、後に説明するようにポリゴン 27 の頂点に相当する。同様に、円筒座標系の  $\theta_2$  平面での切断輪郭線を多角形近似し、L02 を得る。以上の操作を、すべてのボクセルに対応する  $\theta$  平面に対して行なう。

続いて、これら各輪郭線の各頂点を、それぞれ最近接の 3 頂点と接続することで三角パッチ 27 が生成される。このような輪郭線の多角形近似と最近接 3 頂点の接続という手続で三角パッチ 27 を生成することにより、表現に必要なポリゴン数を少なく抑え、かつ、表現形状の精度を維持することが可能となる。

図 11 は、対象物体 100 をポリゴン 27 を用いて表現した 3 次元形状モデル 300 を示す図である。

以上の操作により、対象物体の形状について、計算機内等で再構成を行なうことが可能となる。

以上、円筒ボクセル空間 251 を仮定して説明したが、直交ボクセル空間を用いても良い。また、ポリゴン生成処理においては、隣接ボクセルを接続することにより、細かいポリゴンを一旦生成した後に、これらを統合してポリゴン数を減らしても良い。

## 20 (テクスチャマッピング)

続いて、以上のようにして、計算機内等において再構成された物体形状に対しテクスチャ情報を付与して、よりリアルな 3 次元モデルとするために、撮影した物体画像 A1 ~ An に含まれるテクスチャ情報を上記 3 次元形状モデル 300 に対して付与する処理を行なう。

25 具体的には、各ポリゴン 27 のテクスチャ情報を与える物体画像（以後、参照画像と呼ぶ）を決定した後、ポリゴン 27 を参照画像に投影し、その投影部分のテクスチャ情報を、対応するポリゴン 27 に付与する。

図 12 A および図 12 B は、このようなテクスチャ情報の付与処理を説明するための概念図である。以下では、説明を簡単とするために、参照画像としては、

ラベル番号が1～8までの8枚の物体画像情報があるものとする。すなわち、対象物体を45°おきの角度から撮影した物体画像が存在している。なお、以下の説明では、対象物体を1つの回転軸のまわりに一定の角度ごとに撮影した参照画像に基づいて、テクスチャ情報を形状モデル300に付与する場合について説明するが、本発明はこのような場合に限定されることなく、任意の位置および方向から対象物体を撮影した複数の参照画像に基づいて、形状モデル300に対してテクスチャ情報を付与する場合にも適用することが可能である。

さて、注目するポリゴン27に対して、どの参照画像を対応付けるかを決定する場合に、考慮すべき点としては、まず当該ポリゴン27に対するテクスチャ情報量の大きい参照画像を選択するという方針が挙げられる。

したがって、上記方針の下に、各ポリゴン27に対して対応する参照画像すなわちラベル番号の割当を行なうことで、ポリゴン27で表現された形状モデル300に対してテクスチャ情報を付与することが可能となる。

図13は、テクスチャ情報を付与するまでの処理の流れを示すフローチャートである。一方、図4は、コンピュータ130中において、テクスチャ情報付与を行なうための色彩情報付与処理部200の構成を示す概略ブロック図である。

色彩情報付与処理部200は、カメラ120により撮影された物体画像情報（参照画像情報）を記憶保持する画像記憶部220と、画像記憶部220に記憶された参照画像情報に基づいて、対象物体の形状モデル300を生成する演算部210と、演算部210により生成された形状モデル300、より具体的には各ポリゴン27の位置および形状を記憶している形状記憶部230と、後に説明するように、画像記憶部220に保持された参照画像情報に基づいて、演算部210が、各ポリゴン27に対して割当てたテクスチャ情報を記憶保持する色彩情報記憶部240を備える。

図13および図4を参照して、前述したとおり、たとえば、対象物体を一定角度ごとに回転させながら撮影された画像情報を、画像記憶部220が記憶する（ステップS20）。

続いて、撮影された画像情報に基づいて、演算部210が、形状モデル300を生成し、その形状データを形状記憶部230に保持させる（ステップS22）。

続いて、以下に説明する手順で、立体形状構成要素（たとえばポリゴン 27）と画像記憶部 220 中に保持された参照画像情報との対応付けを演算部 210 が行なう（ステップ S 24）。

5 演算部 210 は、対応付けられた各ポリゴン 27 ごとのテクスチャ情報を色彩情報記憶部 240 に保持させる（ステップ S 26）。

次に、上記ステップ S 24 における立体形状構成要素と参照画像情報との対応付けの処理についてさらに詳しく説明する。

図 14 は、ステップ S 24 における立体形状構成要素と参照画像情報との対応付けを行なうためのフローをより詳しく示すフローチャートである。

10 以下の処理では、各立体形状構成要素（ポリゴン 27）の法線ベクトルと参照画像を撮影した方向と平行な画像撮影面の法線ベクトルとの方向の一致性により、テクスチャ情報量の大きさを判断する。すなわち、対応するポリゴン 27 に対して最も正対する参照画像が、当該ポリゴン 27 に対するテクスチャ情報が最も大きいと判断を行なう。

15 画面撮影（ステップ S 20）および形状モデル作成（ステップ S 22）の各処理が終了した後、演算部 210 は、まず以下の演算処理で用いる変数の初期化を行なう。

すなわち、立体形状構成要素数を変数  $E_{max}$  に代入し、撮影した画像数を変数  $I_{max}$  に代入する。また以下の処理において、撮影画像に対応するラベル番号をカウントするための補助変数  $I_{cnt}$  を値 0 に初期化する。

さらに、 $i$  番目（ $i = 0 \sim E_{max} - 1$ ）の立体形状構成要素にそれぞれ対応する 1 次元配列変数  $Prod[i]$  の値をすべて値 0 に初期化し、一方、 $i$  番目の立体形状構成要素に対応する参照画像のラベル番号が代入される 1 次元配列変数  $Id[i]$  が、すべて値 -1 に初期化される（ステップ S 2402）。

25 続いて、ポリゴン 27 の番号をカウントするための補助変数  $E_{cnt}$  を値 0 に初期化する（ステップ S 2403）。

続いて、 $I_{cnt}$  番目の画像撮影面の法線ベクトルと  $E_{cnt}$  番目の立体形状構成要素の法線ベクトルとの内積の値が計算され、変数  $V_{tmp}$  に代入される（ステップ S 2404）。

続いて、演算部210は、Ecnt番目の変数Prod[Ecnt]の値と上記変数Vtmpの値とを比較する。

変数Prod[Ecnt]の値が変数Vtmpの値以下であると判断すると（ステップS2406）、変数Prod[Ecnt]に変数Vtmpの値が代入され、同時に変数Id[Ecnt]にその時点でのカウント変数Icntの値が代入される（ステップS2408）。  
5

一方、変数Prod[Ecnt]の値が、変数Vtmpの値よりも大きいと判断すると（ステップS2406）、変数Ecntの値が1だけインクリメントされる（ステップS2410）。

10 続いて、カウント変数Ecntの値が、立体形状構成要素数Emaxよりも小さいと判断されると（ステップS2412）、再びステップS2404に処理が戻り、同一の処理が、次の立体形状構成要素に対して繰返される。

一方、変数Ecntの値が立体形状構成要素数Emaxよりも小さいと判断されると（ステップS2412）、カウント変数Icntの値が1インクリメントされる（ステップS2414）。  
15

続いて、カウント変数Icntの値が撮影画像数Imax以上であるか否かの判断が行なわれる（ステップS2416）。

変数Icntの値が撮影画像数Imaxよりも小さいと判断された場合（ステップS2416）、次の参照画像について、ステップS2403からステップS2412までの処理が繰返されることになる。  
20

一方、変数Icntの値が撮影画像数Imax以上であると判断されると（ステップS2416）、次に処理が移行する。

すなわち、ステップS2402からステップS2416までの処理により、各参照画像について、当該参照画像の面法線ベクトルとすべての立体形状構成要素の面法線ベクトルとの内積の値が比較される。この処理により、それまでに処理  
25 の対象となった参照画像についての内積値よりも大きな内積値を有する参照画像については、その内積値が立体形状構成要素の各々について、1次元配列変数Prod[Ecnt]に格納される。一方、その時点における参照画像のラベル番号が1次元配列変数Id[Ecnt]に格納される。



したがって、ステップS 2 4 1 6の処理から次の処理に移行する時点では、1次元配列変数I d [ i ]には、対応するi番目の立体形状構成要素に対して、最大の内積値を有する参照画像情報のラベル番号が格納されていることになる。

5 続いて、色彩情報記憶部2 4 0に対して、演算部2 1 0は、各立体形状構成要素について、対応する参照画像情報を画像記憶部2 2 0から読出して、色彩情報記憶部2 4 0に格納する（ステップS 2 4 1 8）。

10 以上のようにして、形状モデル3 0 0を構成する各立体形状構成要素（ポリゴン2 7）に最もテクスチャ情報量の多い参照画像情報から得た色彩情報（テクスチャ情報）を付与する構成となっているので、最も現実の物体に近いテクスチャ情報を各立体形状構成要素に付与することが可能となる。

図1 5は、図1 4に示したテクスチャ付与方法をコンピュータ1 3 0に実行させるためのプログラムを記録した記録媒体の構成を示す概念図である。

15 記録媒体としては、たとえば磁気ディスクやCD-ROM（コンパクトディスクーリードオンリメモリ）などの光磁気ディスクなどを用いることが可能である。記録媒体2 6 0中には、図1 4に示した処理をコンピュータ1 3 0に実行させるために、所定のプログラム言語により、各処理ステップが記述されたものが、符号化されて記録されている。

20 以上のようにして、記録媒体2 6 0中に記録されたテクスチャ情報付与プログラムに基づいて、コンピュータ1 3 0を動作させることで、上述したのと同様な効果、すなわちコンピュータ1 3 0内で再構成された形状モデル3 0 0に対して、より現実の物体が有するテクスチャに近いテクスチャ情報を付与することが可能となる。

### 〔実施の形態2〕

25 実施の形態1では各立体形状構成要素に対して、内積値に基づいて最もテクスチャ情報量の大きいと判断される参照画像を選択して、各立体形状構成要素にテクスチャ情報を付与した。

しかしながら、図2において説明したとおり、対象物体の形状によっては、ある方向から撮影した物体画像情報には見通せない部分が存在する場合がある。この場合は、この見通せない部分に相当する立体形状構成要素の面法線ベクトルと

最も大きな内積値を有する参照画像には、テクスチャ情報が全くないという場合が存在し得る。

- 実施の形態2では、このような場合にも適用することが可能なテクスチャ情報の付与方法およびテクスチャ情報付与装置ならびにテクスチャ情報付与プログラムを記録した媒体を提供する。

実施の形態2においても、色彩情報付与処理部の構成は、図4に示した色彩情報付与処理部200の構成と同等である。以下に説明するとおり、演算部210の行なう演算処理が、実施の形態1とは異なる。

- すなわち、実施の形態1では、各ポリゴン27について、その法線ベクトルと各参照画面の法線ベクトルとの内積値を比較することで、テクスチャ情報量の大きさを判断した。これに対し、実施の形態2では、各ポリゴン27について、参照画像に対する投影面積を算出し、この算出された投影面積に基づいて、各参照画面情報が有するテクスチャ情報量の大きさを評価する。

- 図16は、このような参照画像へのポリゴン27の投影面積に基づいて、各ポリゴン27ごとに対応する参照画像情報のラベル番号を決定する処理のフローを示すフローチャートである。

- 図16に示すフローチャートでは、評価の対象となる値が、参照画像面の法線ベクトルと立体形状構成要素の法線ベクトルとの内積値 $V_{tmp}$ の代わりに、参照画像に投影された立体形状構成要素の投影面積 $A_{tmp}$ となっている点を除いては、図14に示したフローチャートと全く同様である。

- したがって、ステップS2422からステップS2436までの処理が終了した時点で、1次元配列変数 $Id[i]$ には、対応するi番目の立体形状構成要素が最も大きな投影面積を有する参照画像情報のラベル番号が、1次元配列変数 $Area[i]$ には、対応するi番目の立体形状構成要素について、 $Id[i]$ のラベル番号を有する参照画像情報に対する投影面積が、それぞれ格納されていることになる。

それに応じて、演算部210は、色彩情報記憶部240へ、各立体形状構成要素ごとに、対応する参照画像のテクスチャ情報を画像記憶部220から読出して格納することになる。

以上の処理により、比較的複雑な形状を有する対象物体に対しても、計算機内等において再構成された形状モデル300に対して最もテクスチャ情報量の大きい参照画像情報から各ポリゴン27へのテクスチャ情報を付与することが可能となる。また、図16に示したようなステップS2422からステップS2438  
5 までのプログラムを記録した媒体によりコンピュータ130を動作させることで、同様の効果が奏される。

### [実施の形態3]

以上の説明では、対象物体100を形状モデル300として再構成した場合に、各ポリゴン27に対する参照画像の対応付けは、当該ポリゴン27に対するテク  
10 スチャ情報量の大きさに基づいて判断された。

しかしながら、各ポリゴン27に対して適切な参照画像を決定する場合において考慮すべき点は、テクスチャ情報量の大きさのみに限られない。たとえば、さらにポリゴン27間で付与されたテクスチャ情報に著しい不連続が存在すると、ポリゴン境界線が目立ってしまい、再構成された立体モデル29は極めて不自然  
15 なものとなる。

したがって、実施の形態3における各立体形状構成要素への参照画像の割当方法、すなわちテクスチャ情報付与方法としては、第1にテクスチャ情報量の大きい参照画像を選択すること、および第2にポリゴン境界線を隠蔽することを同時に可能とすることを目的とする。

20 ここで、実施の形態2において説明したように、テクスチャ情報量の大きい参照画像を選択するためには、対応する参照画像へのポリゴンの投影面積は大きい方が望ましい。

一方、ポリゴン境界線を隠蔽するためには、隣接するポリゴン27間に付与される色彩情報（テクスチャ情報）の連続性が高いことが望ましい。

25 実施の形態3では、ポリゴン境界線を隠蔽するために、注目するポリゴン27に対して、それに隣接するポリゴン27への参照画像の割当は、同一であるか、異なる場合でもなるべく撮影角度差の小さい参照画像を選択する構成としている。

より具体的には、上述した2つの条件を最適に満たすポリゴン27への参照画像情報の割付けを可能とするために、以下に説明するように、問題をいわゆるエ

エネルギー最小化問題として扱う。

- すなわち、各参照画像は、一定角度ずつ撮影角度を変化させて撮影されているため、順番に参照画像に番号付けをしておき、各ポリゴン27と参照画像番号との対応付け（ラベリング）問題を以下の式で表わされるエネルギーを局所的に最小化するよう反復改善処理で解くというものである。

なお、各参照画像が一定角度ずつ撮影角度を変化させて撮影されていない場合、すなわち、撮影角度の変化量がそれぞれ異なっている場合、上記番号付けは、撮影角度に対応させれば良い。

$$E = \sum_i \text{Penalty}(i) - k \times \text{area}(i) \quad \dots (1)$$

- ここで、 $\text{Area}(i)$  はポリゴン  $i$  の参照画像への投影面積を表わし、 $\text{Penalty}(i)$  はポリゴン  $i$  とその隣接ポリゴンとの参照画像番号（ラベル）の差を表わし、 $k$  はこれらの結合係数である。

- すなわち、エネルギー関数  $E$  は、ポリゴン  $i$  に隣接するポリゴンに割当てられた参照画像番号が、ポリゴン  $i$  に割当てられた参照画像番号と異なるほど増加し、ポリゴン  $i$  の参照画像への投影面積、すなわちテクスチャ情報量が増加するほど減少する関数である。

- 上述したとおり、ポリゴン  $i$  と隣接するポリゴンに割当てられる参照画像番号の差が小さいほどテクスチャ連続性が高く、ポリゴン境界線が隠蔽されるので、関数  $E$  を極小化することは、テクスチャ情報量の大きさ（色彩情報量の大きさ）とテクスチャ連続性の双方を考慮した場合に、各ポリゴンへの最適な参照画像番号の割当に対応することになる。

- なお、上記のエネルギー関数においては、テクスチャ情報量の大きさとして、ポリゴン  $i$  の参照画像への投影面積を用いたが、実施の形態1に示したように、ポリゴンの面法線ベクトルと参照画像の面法線ベクトルとの内積値で評価する構成とすることも可能である。

ここでエネルギー関数  $E$  において、結合定数  $k$  は定数でも良いし、各ポリゴンの関数（たとえば各ポリゴンの面積の関数）でもよい。

さらに、エネルギー関数  $E$  としては、上述のような関数  $\text{Penalty}(i)$  と

関数  $A r e a (i)$  との線形結合以外にも、より一般的に、注目するポリゴン  $i$  とそれに隣接するポリゴンへ割当てられたテクスチャ情報の連続性の向上に伴い減少し、かつ、注目するポリゴン  $i$  へのテクスチャ情報量の増加に伴い減少する関数であれば、上記の構成には限定されない。

- 5      図 17 は、上記のようなエネルギー関数  $E$  について、反復改善処理で最適値を求めるための処理を示すフローチャートである。

まず、生成された形状モデル 300 の各ポリゴンに対して、初期設定として参照画像番号の仮の対応付けを行なう（ステップ S 2440）。

- 次に、演算部 210 は、変数  $N$  に立体形状構成要素数を代入し、カウント変数  
10       $C n t$  を数値 0 に初期化する。また、フラグ変数  $F l g$  に “OFF” を代入する（ステップ S 2442）。

続いて、変数  $P r e\_l b l$  に  $C n t$  番目の立体形状構成要素に対応する参照画像番号を代入しておく（ステップ S 2444）。

- 続いて、 $C n t$  番目の立体形状構成要素について、対応する参照画像番号を変  
15      化させ、エネルギー関数  $E$  を極小化する参照画像番号を導出する（ステップ S 2446）。

ここで、ステップ S 2446 において得られた新しい対応参照画像番号を変数  $N e w\_l b l$  に代入する（ステップ S 2448）。

- 続いて、変数  $N e w\_l b l$  と変数  $P r e\_l b l$  との値が比較される。両者  
20      が等しくない場合は（ステップ S 2450）、エネルギー関数  $E$  の最小化計算により、ラベルの付け換えが行なわれたと判断され、フラグ変数  $F l g$  に “ON” が代入される（ステップ S 2452）。続いて、カウント変数  $C n t$  の値が 1 だけインクリメントされる（ステップ S 2454）。

- 一方、変数  $N e w\_l b l$  と変数  $P r e\_l b l$  の値が等しい場合は、フラグ  
25      変数は変化されずに、カウント変数  $C n t$  の値のみが 1 だけインクリメントされる（ステップ S 2454）。

続いて、カウント変数  $C n t$  の値が、立体形状構成要素数  $N$  よりも小さい場合は、再び処理がステップ S 2444 に復帰する。一方、カウント変数  $C n t$  の値が立体形状構成要素数  $N$  以上である場合は次に処理が移行する（ステップ S 24

56)。

したがって、ステップS 2 4 4 4からステップS 2 4 5 4までの処理が、すべての立体形状構成要素について繰返されることになる。

続いて、フラグ変数F l gと“O F F”との比較が行なわれる。フラグ変数F l gが“O F F”と等しくない場合は(ステップS 2 4 5 8)、少なくとも1回以上エネルギー関数Eの極小化計算によりラベルの付け換えが行なわれた、すなわちまだエネルギー関数Eを局所的に極小化するラベル番号の対応付けが完了していないと判断され、処理は再びステップS 2 4 4 2に復帰する。

一方、フラグ変数F l gが“O F F”と等しい場合は、ステップS 2 4 4 4からステップS 2 4 5 6までの処理において、エネルギー関数Eを極小化する演算を行なっても、ラベルの付け換えが行なわれていないことを意味する。これは、すなわち現状のラベル番号の対応付けがエネルギー関数Eを局所的に極小化する対応付けとなっていることを意味するため、最適な対応付けが完了したものとして処理が終了する(ステップS 2 4 6 0)。

15 以上の処理により、複数のポリゴンに対して、参照画像番号の対応付けを行なうにあたり、各ポリゴンに対しテクスチャ情報量の大きい参照画像情報を選択するという条件と、ポリゴン境界線を隠蔽するという条件の2つの条件を同時に最適化するテクスチャ情報の割当が行なわれることになる。

したがって、割当が完了した後の形状モデル3 0 0は、現実の物体により近い色彩を有し、かつより自然なテクスチャ連続性を有していることになる。

20 上述のステップS 2 4 4 0からステップS 2 4 6 0までのプログラムを記録した媒体によりコンピュータ1 3 0を動作させることでも、同様の効果を得ることができる。

25 なお、繰返し改善処理の順序は、改善処理の最終結果に影響を与えるため、適切な処理順序を考慮した方が望ましい。これは、繰返し改善処理においては、各ポリゴンのラベル番号を改善する際、その隣接ポリゴンのラベル番号を正しい、あるいは信頼性が高いと仮定しているためである。したがって、より信頼性の低いポリゴンから順に改善処理を行なえば、より良好な改善結果が得られる。

ポリゴンの信頼性を評価する方法の一つとして、ポリゴンの面積や、ポリゴン

が参照画面に投影された際の面積などがある。

これは、上述したステップ S 2 2 4 0 中で行なわれる参照画像番号の仮の対応づけ処理の信頼性が、小さい面積のポリゴンほど、あるいはポリゴンが参照画面に投影された際の面積が小さいポリゴンほど低いためである。

5        [実施の形態 4]

実施の形態 3 のテクスチャ情報付与方法では、テクスチャ情報量（色彩情報量）と、ポリゴン境界線の隠蔽、すなわちテクスチャ連続性の双方を考慮して、各ポリゴンにテクスチャ情報を付与する方法であった。

10        しかしながら、実在の物体から画像情報を撮影する場合には、照明等の関係により、特定方向から撮影した画像情報が、それに近接する方向から撮影した画像情報に対してさえも、光沢等の点で著しく異なっている場合が存在し得る。

したがって、よりテクスチャ連続性の高く、ポリゴン境界線が目立たないテクスチャ情報を付与するためには、実施の形態 3 の方法では不十分な場合がある。

15        そこで、実施の形態 4 のテクスチャ情報付与方法では、1 つのポリゴンに対して、1 つの参照画像情報からテクスチャ情報を付与するのではなく、複数の参照画像情報から、すなわち複数の方向から撮影した画像情報に基づいて、対応するポリゴンにテクスチャ情報を付与する方法を提供する。

実施の形態 4 のテクスチャ情報付与方法の説明をする前に、まず色彩情報記憶部 2 4 0 へのテクスチャ情報の格納方法についてより詳しく説明しておく。

20        図 1 8 は、色彩情報記憶部へのデータの格納方法を示す概念図である。

色彩情報記憶部 2 4 0 は、立体形状構成要素の基本形状、およびテクスチャを記憶するものである。ここで、参照画像情報上の立体形状構成要素は、投影された形状となっているため本来の形状とは異なった形となっている。

25        このため、形状変換を行なって色彩情報記憶部 2 4 0 に色彩を格納する必要がある。

ここで、たとえば立体形状構成要素が 3 角形の場合を例にとってその形状変換を説明する。今、基本形状のテクスチャ情報を図 1 8 で示すような 2 次元離散空間で記憶する場合を考える。基本形状の頂点を  $(x_0, y_0)$  ,  $(x_1, y_1)$  ,  $(x_2, y_2)$  、参照画像情報に対して投影された立体形状構成要素の頂点を

(X 0, Y 0), (X 1, Y 1), (X 2, Y 2) とする。これらを、以下に示す変換行列 A および平行移動ベクトル B により 1 次変換することで、投影された三角形形状を本来の形状に変換することが可能となる。

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad B = \begin{pmatrix} e \\ f \end{pmatrix} \quad \cdots (2)$$

- 5      この場合、基本形状の画素 (x<sub>n</sub>, y<sub>n</sub>) のテクスチャ情報は、以下の式で計算される参照画像情報上の画素 (X<sub>n</sub>, Y<sub>n</sub>) から獲得すればよいことを意味する。

$$\begin{pmatrix} X_n \\ Y_n \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_n \\ y_n \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} \quad \cdots (3)$$

- すなわち上記のようないわゆるアフィン変換により、投影された三角形形状のポリゴンに対して、本来のポリゴン形状に対するテクスチャ情報を獲得し、色彩情報記憶部 240 に格納する。

なお、以上の説明では、ポリゴンの形状を三角形としたが、たとえばその形状が四角形や他の形状の場合も同様に計算することが可能である。

- また、座標変換の方法はアフィン変換の他、射影変換を用いてもよい。射影変換は以下の式で計算される。

$$X_n = \frac{a_1 x_n + a_2 y_n + a_3}{a_7 x_n + a_8 y_n + 1}, \quad Y_n = \frac{a_4 x_n + a_5 y_n + a_6}{a_7 x_n + a_8 y_n + 1} \quad \cdots (4)$$

以上説明したとおり、色彩情報記憶部 240 には、参照画像情報上に投影されるポリゴンの形状にかかわらず、本来のポリゴン形状に対応したテクスチャ情報が格納されている。

- 20      そこで、たとえば実施の形態 3 に示したようなエネルギー関数 E に対する反復改善処理により、各ポリゴン i に対して対応する参照画像情報番号の割当が完了したものとす。

- 実施の形態 4 のテクスチャ情報付与方法では、このラベル番号の割当の完了後に、さらに以下に説明するような重み平均処理を行なうことで、テクスチャ連続性をより向上させるテクスチャ情報の付与方法を実現することを目的とする。

図 19 は、上述したように、各ポリゴン i に対して、参照画像情報番号の割付



が完了した後に行なう重み平均処理のフローを示すフローチャートである。

したがって、たとえば図 17 に示したフローにおけるステップ S 2 4 6 0 に続く処理を示している。

5 まず、初期設定として、変数 E m a x に立体形状構成要素数が、変数 I m a x に撮影された参照画像情報数がそれぞれ代入される。一方、カウンタ変数 E c n t の値が 0 に初期設定される (ステップ S 2 5 0 0)。

続いて、カウンタ変数 I c n t ならびに変数 w a c c の値が 0 に初期設定される (ステップ S 2 5 0 1)。

10 続いて、I c n t 番目の参照画像情報が、E c n t 番目の立体形状構成要素のテクスチャ情報の入力対象であるか否かの判断が行なわれる (ステップ S 2 5 0 2)。

15 ここで、I c n t 番目の画像が、テクスチャ情報の入力対象であるとは、既に行なわれた、ポリゴン (立体形状構成要素) に対する参照画像番号の割当により割当てられた画像情報のみならず、それに隣接する所定数、たとえば前後 1 画面の参照画像情報も入力対象に含めることを意味する。

次に、変数 w g h t に対して、I c n t 番目の参照画像情報に投影された E c n t 番目の立体形状構成要素の面積の値が代入される (ステップ S 2 5 0 4)。

20 続いて、E c n t 番目の立体形状構成要素のテクスチャ情報として、I c n t 番目の参照画像情報の情報を変数 w g h t だけ重み付けをした後、色彩情報記憶部 2 4 0 に格納する (ステップ S 2 5 0 6)。

変数 w a c c に対して、変数 w g h t の値を累積し (ステップ S 2 5 0 8)、カウンタ変数 I c n t の値が 1 だけインクリメントされる (ステップ S 2 5 0 9)。

25 続いて、カウンタ変数 I c n t の値が撮影された参照画像数 I m a x と比較される (ステップ S 2 5 1 0)。

変数 I c n t が、変数 I m a x よりも小さい場合は、処理が、再びステップ S 2 5 0 2 に復帰する。

一方、ステップ S 2 5 0 2 において、I c n t 番目の参照画像情報が、E c n t 番目の立体形状構成要素のテクスチャの入力対象でないと判断された場合は、

処理がステップS 2 5 0 9に移行し、変数I c n tの値が1だけインクリメントされた後（ステップS 2 5 0 9）、変数I c n tと変数I m a xの値の比較が行なわれる。

したがって、ステップS 2 5 0 0からステップS 2 5 1 0までの処理が繰返されることで、E c n t番目の立体形状構成要素に対して、所定数の参照画像情報から重み付けされたテクスチャ情報が獲得され、そのテクスチャ情報が色彩情報記憶部2 4 0に累積されることになる。

続いて、色彩情報記憶部2 4 0に累積されたテクスチャ情報が変数w a c cの値で割算される（ステップS 2 5 1 2）。

ここまでの処理で、E c n t番目の立体形状構成要素に対するテクスチャ情報が、対応する所定数の参照画像情報からのテクスチャ情報の重み付け平均として色彩情報記憶部2 4 0に格納されることになる。

以上の処理を言い換えれば、各ポリゴンについて、それぞれに割り当てられた参照画像番号に対応する参照画像情報、および、これに隣接する所定数の物体画像情報に投影されたポリゴンの面積を求め、これを下記重み平均処理を行なう際の重み係数とする。

ここで、前記物体画像情報に番号を仮定し、これをI c n tとしたとき、この物体画像情報に対応する重み係数をw g h t (I c n t)とする。また、これら画像情報の数をNとする。

さて、ポリゴンのテクスチャ情報は、複数の画素から構成されており、今、1つのテクスチャ情報の画素に着目する。この画素が上記物体画像情報に投影される部位を求め、その投影部位の画像情報（投影部位の画素値、すなわち色、濃度あるいは輝度）を、上記物体画像情報全てにわたって、すなわちN個分について重み平均処理をすることで、いま着目したテクスチャ情報の画素値の値とする。

ここで、投影された部位の画像情報がv (I c n t)であったとすると、重み平均処理は、具体的には以下の式で表される計算を行なうことに対応する。

$$\left[ \sum \text{wght}(\text{Icnt}) \times v(\text{Icnt}) \right] / \sum \text{wght}(\text{Icnt}) \quad \dots (5)$$

この処理を、ポリゴンのテクスチャ情報となるすべての画素について行なう。

続いて、変数E c n tの値が1だけインクリメントされる（ステップS 2 5 1

4)。

続いて、カウント変数  $E_{cnt}$  の値と立体形状構成要素数  $E_{max}$  との値が比較される (ステップ S 2 5 1 6)。

変数  $E_{cnt}$  の値が、立体形状構成要素数  $E_{max}$  よりも小さい場合は、処理がステップ S 2 5 0 1 に復帰する。したがって、すべての立体形状構成要素について、上述したテクスチャ情報の重み平均処理が行なわれることになる。

一方、カウント変数  $E_{cnt}$  が、立体形状構成要素数  $E_{max}$  以上である場合は (ステップ S 2 5 1 6)、色彩情報記憶部 2 4 0 へのテクスチャ情報の格納処理が終了する (ステップ 2 5 1 8)。

すなわち、実施の形態 4 のテクスチャ情報の付与方法では、まず、各ポリゴンに対して参照画像情報番号 (ラベル番号) の対応付けが行なわれる。続いて、対応付けられた参照画像情報番号を含む所定数 (たとえば、対応付けられた参照画像情報番号とその前後 1 画面分) の参照画像情報番号に対応する参照画像情報について、各参照画像情報に投影される立体形状構成要素の面積に応じた重み平均処理を行なった結果が、その立体形状構成要素のテクスチャ情報として付与される。

以上説明したような所定数の参照画像情報からのテクスチャ情報の重み平均処理を行なうことにより、対応するポリゴンに対するテクスチャ情報が得られるので、よりテクスチャ連続性の向上したテクスチャ情報を当該ポリゴンに割当てることが可能となる。

たとえば、現実の対象物体を撮影した場合に、照明等の関係により、特定方向から撮影した参照画像情報における当該ポリゴンに対する色彩情報に含まれる光沢が特異的に高い場合でも、重み平均処理によりその影響を軽減することが可能である。

図 1 9 に示したようなステップ S 2 5 0 0 からステップ S 2 5 1 8 までのプログラムを記録した媒体により、コンピュータ 1 3 0 を動作させることでも、同様の効果を得ることが可能である。

[実施の形態 5]

実施の形態 4 では、予め、各ポリゴンに対して、テクスチャ情報を獲得する参

照画像番号の割当が完了した後に、所定数の隣接する参照画像情報からも対応するポリゴンに対してテクスチャ情報を与える構成について説明した。

- 5       しかしながら、テクスチャ連続性をより重視する考えに立てば、必ずしもエネルギー関数Eを極小化するように各ポリゴンに対する参照画像番号の割当を行なう必要はない。

実施の形態5のテクスチャ情報付与方法では、各ポリゴン（立体形状構成要素）に対して、その立体形状構成要素に対するテクスチャ情報を有する複数の参照画像情報から当該ポリゴンに対してテクスチャ情報を付与する構成としている。

- 10       たとえば、各ポリゴン（立体形状構成要素）に対して、その立体形状構成要素に対するテクスチャ情報を有するすべての参照画像情報から当該ポリゴンに対してテクスチャ情報を付与しても良いし、各ポリゴン（立体形状構成要素）に対して、その立体形状構成要素に対するテクスチャ情報を有する画像情報からランダムに、あるいは規則的に参照画像情報を選択し、これらから当該ポリゴンに対してテクスチャ情報を付与しても良い。

- 15       図20は、このようなポリゴンへのテクスチャ情報の付与方法を示す概念図であり、立体形状構成要素に対するテクスチャ情報を有するすべての参照画像情報から当該ポリゴンに対してテクスチャ情報を付与する例を示している。

- 20       実施の形態4において説明したとおり、色彩情報記憶部240に対しては、各参照画像情報に投影されるポリゴンの形状にかかわらず、本来のポリゴン形状に対応したテクスチャ情報が格納される。

したがって、特定のポリゴンiに注目した場合、このポリゴンが投影される面積が0でないすべての参照画像情報から、その投影面積に応じた重み平均処理を行なうことによりテクスチャ情報を獲得することが可能である。

- 25       図21は、このようなテクスチャ情報付与方法のフローを示すフローチャートである。

実在の物体に対する複数の画像撮影（ステップS20および形状モデル作成ステップS22）の後、各立体形状構成要素と、当該立体形状構成要素の投影面積が0でない参照画像情報との対応付けを行なう（ステップS30）。

続いて、上記対応付けに基づいて、色彩情報記憶部240に対して、投影面積

に応じた重み平均処理を行なうことにより、各立体形状構成要素ごとに、テクスチャ情報が累積される（ステップS 3 2）。

つまり、実施の形態5のテクスチャ情報の付与方法では、立体形状構成要素ごとに、複数の参照画像情報のそれぞれに投影された当該立体形状構成要素の面積を重み係数として、複数の参照画像情報についての重み平均処理が行なわれる。この立体形状構成要素ごとに得られた重み平均処理の結果が、テクスチャ情報として各立体形状構成要素に付与される。

以上のようにして、各立体形状構成要素に対して、テクスチャ情報を有するすべての参照画像情報から、テクスチャ情報を付与する構成とすることで、テクスチャ連続性がさらに向上する。

つまり、特定方向から撮影した参照画像情報が、たとえば照明の関係等で他の方向から撮影した参照画像情報に比べて光沢が異常に高いような場合でも、すべての関係する参照画像情報からのテクスチャ情報を重み平均することで、この特定方向のテクスチャ情報の影響が抑制される。

なお、以上説明した実施の形態1～5では、形状モデル300をポリゴンデータに変換してから、テクスチャ情報を付与しているが、本願はこのような構成に限定されるわけではなく、ボクセル表現された形状モデル300において、表面の面方向を演算して、テクスチャ情報を付与するようにしてもよい。

ただし、ポリゴンデータに変換してから、テクスチャ情報を付与するという構成の方が、同一方向に向いている面（ポリゴン）を一括して処理できるため、演算量を著しく減少させることが可能である。

#### 〔実施の形態6〕

図22は、本発明の実施の形態6による物体抽出装置（画像切出装置）の全体構成を示す図である。図22を参照して、この物体抽出装置は、コンピュータ130を備える。コンピュータ130は、CD-ROM260に記録されているプログラム301に従って、物体画像中の物体部分を検出し、抽出する。このプログラム301は、物体画像の領域分割処理を行なうステップS1と、領域情報の記憶処理を行なうステップS2と、各領域の物体画像と背景画像との差分処理を行なうステップS3と、各領域において、差分の絶対値の平均値導出を行なうス

ステップS 4と、差分の絶対値の平均値としきい値との比較による物体部分の検出処理を行なうステップS 5と、検出した物体部分を抽出するステップS 5とを含む。ステップS 1～S 5の詳細については後で詳しく説明する。

図23は、本発明の実施の形態6による物体抽出装置（画像切出装置）の概略を示すブロック図である。図23を参照して、物体抽出装置としてのコンピュータ130は、画像記憶部220、演算部210、領域情報記憶部241および抽出画像記憶部231を備える。なお、各部220、210、231、241の詳細は後述する。

図24は、図23の演算部210の概略を示すブロック図である。図24を参照して、演算部210は、領域分割部9および抽出部10を備える。抽出部10は、差分処理部11、平均値導出部13、しきい値処理部15および物体部分抽出部16を備える。物体画像Aは、カメラなどの撮影装置によって、対象物体を、背景とともに撮影することにより得られる。背景画像Bは、カメラなどの撮影装置によって、対象物体の背景だけを撮影することにより得られる。背景画像Bおよび物体画像Aは、図23の画像記憶部220に記憶される。ここで、撮影対象となる背景は一般に対象物体の後方にあるものであるが、対象物体の前方にあるものもある。

領域分割部9は、物体画像Aを複数の領域に分割する（図22のプログラム301のステップS1）。領域分割に関する情報は、図23の領域情報記憶部241に記憶される（図22のプログラム301のステップS2）。差分処理部11は、領域分割部9によって得られた領域単位で、物体画像Aと背景画像Bとの間で差分処理を行ない、差分を求める（図22のプログラム301のステップS3）。差分は、物体画像Aと背景画像Bの色情報の差を、画素単位で求めたものである。平均値導出部13は、差分の絶対値を求め、領域単位で差分の絶対値の平均値を導出する（図22のプログラム301のステップS4）。つまり、平均値導出部13は、領域ごとの差分の絶対値の平均値を導出する。しきい値処理部15は、各領域における差分の絶対値の平均値と、しきい値とを比較し、差分の絶対値の平均値がしきい値以上の領域を物体部分として検出する（図22のプログラム301のステップS5）。なお、しきい値は、経験に基づいて設定される。

物体部分抽出部 1 6 は、しきい値処理部 1 5 で検出された物体部分を抽出する（図 2 2 のプログラム 3 0 1 のステップ S 6）。つまり、物体部分抽出部 1 6 は、物体画像から、しきい値処理部 1 5 で検出された物体部分を取り出す。抽出された物体部分の画像は、図 2 3 の抽出画像記憶部 2 3 1 に記憶される。

- 5        領域分割部 9 が行なう領域分割について詳しく説明する。領域分割は、たとえば、「高木他編、別冊 O p l u s E、画像処理アルゴリズムの最新動向、p p . 2 2 7 - 2 3 3、新技術コミュニケーションズ」に記されているように、一般に用いられている、エッジ延長法、領域・エッジ併用法、F a c e t m o d e l による領域分割などを用いる。このうちエッジ延長法について説明する。まず第 1 に、画像の 1 次微分から、各画素についてエッジ強度とエッジ方向を計算
- 10        する。第 2 に、エッジ強度についての極大値抑制処理としきい値処理によって、極大かつ一定値以上のエッジ強度を持つエッジ要素（強いエッジ要素と呼んでいる）を抽出する。この段階では、強いエッジ要素は必ずしも連続していない。第 3 に、強いエッジ要素のうち端点になっているものを始点としてエッジの延長を行なう。以上が、エッジ延長法である。
- 15

- 図 2 5 A から図 2 5 C は、図 2 4 の差分処理部 1 1、平均値導出部 1 3、しきい値処理部 1 5 および物体部分抽出部 1 6 における処理を詳しく説明するための図である。図 2 5 A を参照して、物体画像 1 7 は、物体部分 1 9 と背景部分 2 1 とからなる。背景画像 2 3 は、背景 2 5 だけからなる。物体画像 1 7 は、図 2 4
- 20        の領域分割部 9 によって、複数の領域  $a_1 \sim a_n$  に分割されている。

- 領域  $a_1$  に注目して、図 2 4 の差分処理部 1 1 の動作について説明する。図 2 5 B も参照して、領域  $a_1$  の各画素の色情報と、領域  $a_1$  に対応する背景 2 5 の領域  $b_1$  の各画素の色情報との差を求める。このようにして、領域  $a_1$  における差分の集合  $c_1$  を得る。そして、図 2 4 の平均値導出部 1 3 は、差分の集合  $c_1$
- 25        を構成する差分の絶対値を求め、差分の絶対値の平均値を求める。さらに、図 2 4 のしきい値処理部 1 5 は、差分の集合  $c_1$  を構成する差分の絶対値の平均値と、しきい値とを比較し、平均値がしきい値以上の場合は、差分の集合  $c_1$  に対応する領域  $a_1$  を物体部分として検出する。差分処理部 1 1、平均値導出部 1 3 およびしきい値処理部 1 5 は、すべての領域  $a_1 \sim a_n$  について、上述したような差

分処理、差分の絶対値の平均値導出およびしきい値処理を行なう。そして、物体部分抽出部 16 は、しきい値処理部 15 によって検出された物体部分を、物体画像 17 から抽出する。図 25 C は、以上のようにして抽出した物体部分 19 を示す。したがって、背景部分 21 のような不所望部分が除かれている。もちろん、  
5 対象物体の前方に位置する物が物体画像 17 中に含まれている場合は、その部分も不所望部分として除かれる。

以上のように、本発明の実施の形態 6 による物体抽出装置では、物体画像を複数の領域に分割し、領域単位で差分の絶対値の平均値を求め、平均値がしきい値以上の領域を物体部分として抽出する。このため、本発明の実施の形態 6 による  
10 物体抽出装置、物体抽出方法および物体抽出プログラムを記録した媒体では、画素レベルで、対象物体に背景と同色の部分が存在していても、領域レベルにおいて背景と異なる色が存在すれば、物体部分として検出し、抽出できる。また、人手による作業が少なくなるとともに、同一色の背景板を用いなければならないような特殊な撮影環境が不要である。

15 図 24 の差分処理部 11 における差分処理の他の例について説明する。上述の説明では、差分処理部 11 は、領域単位で差分を求めるが、領域単位ではなく、物体画像全体と背景画像全体との間で差分処理を行ない、差分を求めることができる。そして、平均値導出部 13 において、領域分割部 9 で求めた領域単位で、差分の絶対値の平均値を導出する。

20 また、物体画像の各領域内の画素の平均値を算出し、その平均値とその領域に対応する背景画像の領域内の画素の平均値との差の絶対値を算出し、その差の絶対値を予め定められた値と比較し、その差の絶対値が予め定められた値以上の領域を物体部分として抽出してもよい。

上記実施の形態 6 では、領域分割は、エッジを参考に求めたが、本発明は当然、  
25 同色の部分を同一領域とするようにしてもよい。また、複数の領域分割手法を組合せてもよい。

また、上記実施の形態 6 では、カラー画像について述べたが、当然、本発明は白黒画像であってもよく、また、上述した色情報（色信号レベル）に代えて濃度情報（輝度信号レベル）を用いてもよい。



また、上記実施の形態6では、しきい値以上の領域をそのまま、物体部分としたが、本願は1回だけの処理に限定されるわけではない。たとえば、最初の処理で検出した物体部分を仮の物体部分とし、その他を仮の背景部分とする。そして、この物体画像の仮の背景部分の明るさと、この仮の背景部分に対応する背景画像の領域の明るさを比較し、背景画像と入力画像との照明状況の変化を検出する。そして、これにより、物体画像の輝度を一様に修正して、同一処理を再度行なってもよい。

また、上記実施の形態6では、しきい値の値は、一定であるが、これに限定されるわけではない。たとえば、画像の中央部分と周辺部分とでは、しきい値の値を変更するようにしてもよい。また、領域の面積に応じて、しきい値の値を変更するようにしてもよい。また、再処理を行なうのであれば、近傍に物体部分が存在するか否かに応じて変更してもよい。

また、上記実施の形態6では、各領域の差分の絶対値を平均し、これをしきい値と比較しているが、これに限定されるわけではない。たとえば、差分の値のばらつき具合をも参考に、物体部分か否かを判定するようにしてもよい。

また、上記実施の形態6では、物体部分を最終的に抽出処理しているが、本願は、これに限定されるわけではない。たとえば、抽出処理ではなく、物体の有無判定に用いるようにしてもよい。なお、このような有無判定は、ビル監視システムの侵入者検知において使用されている。

## 20 [実施の形態7]

図26は、本発明の実施の形態7による物体抽出装置の全体構成を示すフローチャートである。図26のステップS112～S118は、コンピュータ130に、対象物体を撮影して得られた物体画像から背景部分を除くようにして物体部分を抽出させるためのプログラムであって、CD-ROM260に記録されている。

このプログラムは、ステップS111で得られた物体画像の奥行情報 $d_p(i, j)$ を各画素 $(i, j)$ ごとにステレオ法によって算出するステップS112と、物体画像を複数の領域Rに分割するステップS113と、奥行情報の平均値 $m d_p(R)$ を各領域Rごとに算出するステップS114と、奥行情報の平均値 $m d$

p (R) を予め定められたしきい値 d p t h と比較するステップ S 1 1 5 と、奥行情報の平均値 m d p (R) がしきい値 d p t h よりも大きい場合にその領域 R を背景部分として削除、具体的にはその領域 R 内の各画素の値 v (i, j) を 0 にするステップ S 1 1 6 と、奥行情報の平均値 m d p (R) がしきい値 d p t h よりも小さい場合にその領域 R を物体部分として抽出、具体的にはその領域 R 内の各画素の値 v (i, j) を 1 にするステップ S 1 1 7 と、上記ステップ S 1 1 5 ~ S 1 1 7 の処理を全領域 R について行なったか否かを判定するステップ S 1 1 8 とを含む。ここで、画素の値としては、輝度（濃度）、色情報、またはそれらの組合せなどを用いることができる。

10 次に、この実施の形態 7 による物体抽出装置の動作を図 2 6 のフローチャートを参照して説明する。

まずステップ S 1 1 1 において、デジタルスチルカメラなどを用いて対象物体をその背景とともに撮影し、物体画像を得る。この物体画像はコンピュータ 1 3 0 内の画像記憶部 2 2 0 に格納される。これにより、各画素 (i, j) の値として v (i, j) が得られる。ここでは、静止画を撮影するスチルカメラを用いるが、これに代えて動画を撮影するビデオカメラやデジタルカメラなどを用いてもよい。

続いてステップ S 1 1 2 において、ステレオ法などに従って各画素 (i, j) の奥行情報 d p (i, j) を算出する。ステレオ法は、たとえば D. H. Ballard 他著、「コンピュータビジョン」、Prentice Hall、p p. 8 8 ~ 9 3 に開示されている。ステレオ法によれば、まず所定距離だけ離れた 2 つの視点から対象物体を撮影し、それにより得られた 2 つの物体画像間の対応点を定め、逆透視変換法または単純な三角測量法を用いて奥行情報 d p (i, j) を算出する。ステレオ法の応用例はたとえば特開平 8 - 3 3 1 6 0 7 号公報に開示されている。なお、  
20  
25  
ここでは奥行情報を算出するためにステレオ法を用いたが、動きからの奥行推定方法 (Shape-from-Motion)、類似性および連続性をともに考慮した繰返し改善法 (一種の緩和法) などを用いてもよい。

ステップ S 1 1 2 と並行してステップ S 1 1 3 においては、撮影された物体画像を上記実施の形態 6 と同様に複数の領域 R に分割する。ステップ 1 2 の奥行情

報算出とステップ 1 3 の領域分割とは同時に行なう必要はなく、奥行情報の算出後に領域分割を行なってもよく、逆に、領域分割後に奥行情報の算出を行なってもよい。

図 2 7 A は複数の領域 R に分割された物体画像を示し、図 2 7 B は奥行情報を画素の輝度で表わした画像を示す。輝度の高い画素ほど撮影位置からの距離が近く、輝度の低い画素ほど撮影位置からの距離が遠いことを示す。したがって、物体部分は明るく、背景部分は暗く表わされる。

続いてステップ S 1 1 4 において、各領域 R ごとに次の式 (6) に従って奥行情報の平均値  $mdp(R)$  を算出する。

$$10 \quad mdp(R) = \frac{\sum_R dp(i, j)}{n} \quad \dots (6)$$

$\sum_R$  : 領域 R 内の総和  
 $n$  : 領域 R 内の画素数

続いてステップ S 1 1 5 において、その算出された奥行情報の平均値  $mdp(R)$  がしきい値  $dp_{th}$  と比較される。このしきい値  $dp_{th}$  は経験的に予め定められる。

15 奥行情報の平均値  $mdp(R)$  がしきい値  $dp_{th}$  よりも大きい場合、その領域 R 内の画素の値  $v(i, j)$  がすべて 0 にされる。すなわち、物体画像からその領域 R が背景部分として削除される。他方、奥行情報の平均値  $mdp(R)$  がしきい値  $dp_{th}$  よりも小さい場合、ステップ S 1 1 7 においてその領域 R 内の画素の値  $v(i, j)$  がすべて 1 にされる。すなわち、物体画像からその領域 R  
 20 が物体部分として抽出される。

続いてステップ S 1 1 8 において、上記ステップ S 1 1 5 ~ S 1 1 7 の処理がすべての領域 R について行なわれたか否かが判別される。上記処理がすべての領域 R について行なわれた場合、図 2 7 C に示されるような画像が得られる。

25 以上のようにこの実施の形態 7 によれば、物体画像を複数の領域 R に分割し、各領域 R ごとに奥行情報の平均値を算出し、その平均値が予め定められたしきい値よりも小さい領域 R を物体部分として抽出するため、図 2 7 C に示されるように物体画像から背景部分を除くようにして物体部分だけをその輪郭で正確に切出

することができる。また、奥行情報を用いるため、対象物体の背景だけを別途撮影する必要がない。

〔実施の形態 8〕

図 28 は、本発明の実施の形態 8 による物体抽出装置の主要構成を示すフロー  
5 チャートである。図 28 において、ステップ S 222, S 224 ~ S 227 は、  
コンピュータ 130 に、対象物体を撮影して得られた物体画像と、対象物体の背景  
だけを複数回撮影して得られた複数の背景画像とに基づいて、物体画像から背景  
部分を除くようにして物体部分を抽出させるためのプログラムであって、CD  
-ROM 260 に記録されている。

10 このプログラムは、ステップ S 221 において得られた複数の背景画像において  
互いに同じ座標に位置する画素の平均値  $m(i, j)$  および標準偏差  $\sigma(i, j)$   
を各画素ごとに算出するステップ S 222 と、ステップ S 223 において得  
られた物体画像の各画素の値  $v(i, j)$  とその画素に対応する背景画像の画素  
の平均値  $m(i, j)$  との差の絶対値（以下、単に「差」という。） $|v(i, j) - m(i, j)|$   
15  $|v(i, j) - m(i, j)|$  を算出し、その差  $|v(i, j) - m(i, j)|$  を標準  
偏差  $\sigma(i, j)$  の  $k$  倍と比較するステップ S 224 と、差  $|v(i, j) - m(i, j)|$   
が  $k\sigma(i, j)$  よりも大きい場合にその画素を背景部分として削除、すなわちその画素の値  $v(i, j)$  を 0 にするステップ S 225 と、差  $|v(i, j) - m(i, j)|$   
が  $k\sigma(i, j)$  よりも大きい場合にその画素を物  
20 体部分として抽出、すなわちその画素の値  $v(i, j)$  を 1 にするステップ S 2  
26 と、上記ステップ S 224 ~ S 226 の処理をすべての画素について行なっ  
たか否かを判定するステップ S 227 とを含む。

次に、この実施の形態 8 による物体抽出装置の動作を図 28 を参照して説明する。

25 まずステップ S 221 において、デジタルスチルカメラなどを用いて対象物体  
の背景だけを同じ視点から複数回撮影し、複数の背景画像を得る。ここで、背景  
画像の枚数は精度を考慮すれば 3 枚以上が好ましく、簡便性を考慮すれば 10 枚  
程度が好ましい。

続いてステップ S 222 において、複数の背景画像において互いに同じ座標に

位置する画素の平均値  $m(i, j)$  および標準偏差  $\sigma(i, j)$  がそれぞれ次の式 (7) および (8) に従って各画素ごとに算出される。これにより、画像信号を A/D 変換する A/D 変換器の変換特性および照明特性の変動やジッタが原因で撮影された背景画像の画素値として特異な値が得られても画素値の平均を算出するため、安定した背景画像を得ることができる。

$$m(i, j) = \frac{\sum v(i, j)}{N} \quad \dots (7)$$

$$\sigma(i, j) = \sqrt{\frac{\sum v(i, j)^2}{N} - \left(\frac{\sum v(i, j)}{N}\right)^2} \quad \dots (8)$$

ここで、N は物体画像の全領域 R 内の画素数である。

続いてステップ S 2 2 3 において、対象物体を撮影して物体画像を得る。ここで、物体画像の各画素の値として  $v(i, j)$  が得られる。

続いてステップ S 2 2 4 において、物体画像の各画素の値  $v(i, j)$  とその画素に対応する背景画像の画素の平均値  $m(i, j)$  との差  $|v(i, j) - m(i, j)|$  が算出される。

差  $|v(i, j) - m(i, j)|$  が  $k\sigma(i, j)$  よりも小さい場合、ステップ S 2 2 5 においてその画素の値  $v(i, j)$  が 0 にされる。これにより、その画素は背景部分として物体画像から削除される。他方、差  $|v(i, j) - m(i, j)|$  が  $k\sigma(i, j)$  よりも大きい場合、ステップ S 2 2 6 においてその画素の値  $v(i, j)$  が 1 にされる。これにより、その画素が物体部分として物体画像から抽出される。ここで、k は 3 程度が好ましい。

続いてステップ S 2 2 7 において、上記ステップ S 2 2 4 ～ S 2 2 6 の処理をすべての画素について行なったか否かが判定される。上記処理がすべての画素について行なわれた場合、このプログラムは終了する。

以上のようにこの実施の形態 8 によれば、複数の背景画像に基づいて画素の平均値を算出しているため、画像信号を A/D 変換するための A/D 変換器の変換特性や照明特性の影響が緩和され、しかも物体画像か背景画像かを判別するためのしきい値として複数の背景画像における画素の標準偏差を用いているため、適

切なしきい値を自動的に設定することができる。したがって、物体画像から背景部分だけを除くようにして物体部分を正確に抽出することができる。

[実施の形態 9]

図 29 は、本発明の実施の形態 9 による物体抽出装置の主要構成を示すフロー  
5 チャートである。図 29 において、ステップ S 2 2 2, S 3 3 3 B ~ S 3 3 6,  
S 2 2 7 は、コンピュータ 1 3 0 に、対象物体を撮影して得られた物体画像と、  
対象物体の背景だけを複数回撮影して得られた複数の背景画像とに基づいて、物  
体画像から背景部分を除くようにして物体部分を抽出させるためのプログラムで  
あって、CD-ROM 2 6 0 に記録されている。

10 上記実施の形態 8 ではステップ S 2 2 3 において対象物体を 1 回撮影して 1 枚  
の物体画像を得ているが、この実施の形態 9 ではステップ S 3 3 3 A において対  
象物体を複数回撮影して複数の物体画像を得ている。そのため、このプログラム  
は、複数の物体画像において互いに同じ座標に位置する画素の平均値  $m v(i, j)$ ,  
15  $j)$  を各画素ごとに算出するステップ S 3 3 3 B を含む。ステップ S 3 3 4 ~ 3  
3 6 においては、図 2 8 に示された画素の値  $v(i, j)$  の代わりに画素の平均  
値  $m v(i, j)$  が用いられる。したがって、この実施の形態 9 では、ステップ  
S 3 3 3 A において得られた複数の物体画像において互いに同じ座標に位置する  
画素の平均値  $m v(i, j)$  が各画素ごとに算出される。

続いてステップ S 3 3 4 において、物体画像の各画素の平均値  $m v(i, j)$   
20 とその画素に対応する背景画像の画素の平均値  $m(i, j)$  との差  $|m v(i, j) - m(i, j)|$   
が算出され、その差  $|m v(i, j) - m(i, j)|$  が  $k \sigma(i, j)$  と比較される。

差  $|m v(i, j) - m(i, j)|$  が  $k \sigma(i, j)$  よりも小さい場合、ス  
テップ S 3 3 5 において物体画像のその画素の平均値  $m v(i, j)$  が 0 にされ  
25 る。これにより、その画素が背景部分として削除される。他方、差  $|m v(i, j) - m(i, j)|$   
が  $k \sigma(i, j)$  よりも大きい場合、ステップ S 3 3 6 において物体画像の画素の平均値  $m v(i, j)$  が 1 にされる。これにより、その  
画素が物体画像から物体部分として抽出される。

以上のようにこの実施の形態 9 によれば、対象物体を複数回撮影して得られた

複数の物体画像を用いるため、背景画像と同様に安定した物体画像を得ることができる。したがって、物体画像から背景部分を除くようにして物体部分をより正確に抽出することができる。

[実施の形態 10]

5 図 30 は、本発明の実施の形態 10 による物体抽出装置の主要構成を示すフローチャートである。図 30 において、ステップ S 2 2 2, S 4 4 1 ~ S 4 4 7 は、コンピュータ 1 3 0 に、対象物体を撮影して得られた物体画像と、対象物体の背景だけを複数回撮影して得られた複数の背景画像とに基づいて、物体画像から背景部分を除くようにして物体部分を抽出させるためのプログラムであって、CD-ROM 2 6 0 に記録されている。

図 2 8 に示された実施の形態 5 では物体画像は各画素ごとに処理されるのに対し、この実施の形態 10 では物体画像は複数の領域 R に分割され、各領域 R ごとに処理される。

したがって、このプログラムは、ステップ S 2 2 3 において得られた物体画像を複数の領域 R に分割するステップ S 4 4 1 と、物体画像の各領域 R 内の各画素の値  $v(i, j)$  とその領域 R に対応する背景画像の領域 R 内の対応する画素の平均値  $m(i, j)$  との差を算出し、次の式 (9) で表わされるその差の平均値  $md(R)$  を各領域 R ごとに算出するステップ S 4 4 2 と、ステップ S 2 2 3 において算出された標準偏差の平均値  $m\sigma(R)$  を次の式 (10) に従って各領域 R ごとに算出するステップ S 4 4 3 とを含む。

$$md(R) = \frac{\sum_R |v(i, j) - m(i, j)|}{n} \quad \dots (9)$$

$$m\sigma(R) = \frac{\sum_R \sigma(i, j)}{n} \quad \dots (10)$$

ステップ S 4 4 4 ~ S 4 4 6 においては、図 2 8 に示された差  $|v(i, j) - m(i, j)|$  の代わりに差の平均値  $md(R)$  が用いられ、標準偏差  $\sigma(i, j)$  の代わりに標準偏差の平均値  $m\sigma(R)$  が用いられる。そして、ステップ S 2 2 3 において得られた物体画像はステップ S 4 4 1 において複数の領域 R に分割される。

続いてステップS 4 4 2において、物体画像の各領域R内の各画素の値  $v(i, j)$  とその領域Rに対応する背景画像の領域R内の対応する画素の平均値  $m(i, j)$  との差  $|v(i, j) - m(i, j)|$  が算出され、その差の平均値  $m_d(R)$  が各領域Rごとに算出される。

- 5        続いてステップS 4 4 3において、ステップS 2 2 2において算出された標準偏差  $\sigma(i, j)$  の平均値  $m_\sigma(R)$  が各領域Rごとに算出される。

- 続いてステップS 4 4 4において、差の平均値  $m_d(R)$  が  $k m_\sigma(R)$  と比較される。差の平均値  $m_d(R)$  が  $k m_\sigma(R)$  よりも小さい場合、ステップS 4 4 5においてその領域R内の画素の値  $v(i, j)$  がすべて0にされる。これ  
10        により、その領域Rが背景部分として物体画像から削除される。他方、差の平均値  $m_d(R)$  が  $k m_\sigma(R)$  よりも大きい場合、ステップS 4 4 6においてその領域R内の画素の値  $v(i, j)$  がすべて1にされる。これにより、その領域Rが物体部分として物体画像から抽出される。

- 最後にステップS 4 4 7において、上記ステップS 4 4 4～S 4 4 6の処理を  
15        すべての領域Rについて行なったか否かが判定される。上記処理がすべての領域Rについて行なわれた場合、このプログラムは終了する。

- 以上のようにこの実施の形態10によれば、対象物体を複数の領域Rに分割し、物体画像の各領域R内の各画素の値とその領域Rに対応する背景画像の領域R内の対応する画素の平均値との差の平均値  $m_d(R)$  を各領域Rごとに算出し、その  
20        差の平均値  $m_d(R)$  が標準偏差の平均値  $m_\sigma(R)$  の  $k$  倍よりも大きい領域Rを物体部分として抽出するため、物体画像から背景部分を除くようにして物体部分をより正確に抽出することができる。

- なお、上記実施の形態10のようにステップS 4 4 2において物体画像の各領域R内の各画素の値  $v(i, j)$  とその領域Rに対応する背景画像の領域R内の対応する画素の平均値  $m(i, j)$  との差を算出の方が好ましいが、物体画像  
25        の各領域R内の画素の平均値  $m_v(i, j)$  を算出し、物体画像の各領域R内の画素の平均値とその領域Rに対応する背景画像の領域R内の画素の平均値  $m(i, j)$  との差の絶対値を算出することも可能である。この場合、図30のフローチャートにおいて、物体画像の各領域R内の各画素の値  $v(i, j)$  を物体画像の



各領域R内の画素の平均値 $m_v(i, j)$ に置き換えればよい。

- あるいは、物体画像の各領域R内の画素の平均値 $m_v(R)$ を算出し、また、領域Rに対応する背景画像の領域R内の画素ごとの平均値 $m(i, j)$ の領域R内における平均値 $m_m(R)$ を算出し、さらにこれらの差の絶対値を算出し、この値に基づいて物体部分を抽出することも可能である。この場合、図30のフローチャートにおいて、ステップS442で $m_d(R)$ を求める際に、 $m_d(R)$ として $|m_v(R) - m_m(R)|$ を計算すればよい。

#### [実施の形態11]

- 図31は、本発明の実施の形態11による物体抽出装置の主要構成を示すフローチャートである。上記実施の形態10ではステップS223において対象物体を1回撮影して1枚の物体画像を得ているのに対し、この実施の形態11では上記実施の形態9と同様にステップS333Aにおいて対象物体を同じ視点から複数回撮影して複数の物体画像を得ている。したがって、ステップS551においては複数の物体画像を平均化した物体画像を複数の領域Rに分割している。したがって、ステップS555およびS556においては、画素の値 $v(i, j)$ の代わりに画素の平均値 $m_v(i, j)$ が用いられる。

- この実施の形態11によれば、対象物体を同じ視点から複数回撮影して複数の物体画像を得ているため、対象物体の撮影時におけるA/D変換器の変換特性や照明特性の変動が緩和され、物体画像から背景部分を除くようにして物体部分をより正確に抽出することができる。

#### [実施の形態12]

- 本発明の実施の形態12による立体モデル作成装置は、図3に示された実施の形態1と同様に、回転テーブル110、カメラ120およびコンピュータ130を備える。ここで、回転テーブル110の代わりに、ロボットアームなどを用いることができる。つまり、回転テーブル110の代わりとして、対象物体の向きを変えることができるものを用いることができる。

図32は、この立体モデル作成装置の概略を示すブロック図である。図32を参照して、立体モデル作成装置は、撮影部109、画像記憶部220、演算/制御部113、形状記憶部230および色彩情報記憶部240を備える。撮影部1

09は、図3の回転テーブル110およびカメラ120からなる。画像記憶部220、演算／制御部113、形状記憶部230および色彩情報記憶部240は、図3のコンピュータ130に含まれる。

図33は、図3の立体モデル作成装置における処理の流れを説明するための図である。図6Aから図6Eは、図3の立体モデル作成装置における処理の具体的内容を説明するための図である。図6Aは、図33のステップS2における対象物体・背景の撮影を説明するための図である。図6Bは、図33のステップS12におけるシルエット画像作成を説明するための図である。図6Cは、図33のステップS14におけるボーディング処理を説明するための図である。図6Dは、図33のステップS16におけるポリゴン作成を説明するための図である。図6Eは、図33のステップS18におけるテクスチャマップを説明するための図である。

図3、図6A～図6E、図32、および図33を参照して説明を行なう。ステップS8において、キャリブレーションを行なう。本実施の形態12によるキャリブレーションとは、カメラ120の内部パラメータ（パースペクティブ比）およびカメラ120と回転テーブル110との位置関係を求める処理のことである。ステップS10では、対象物体および背景の撮影を行なう。つまり、回転テーブル110に対象物体を置かずに、背景だけを撮影し、背景画像を1枚得る。また、回転テーブル110に対象物体100を置き、対象物体100を回転させる。そして、カメラ120によって、所定の角度ごとに、対象物体100を背景とともに撮影し、物体画像A1～Anを得る。たとえば、対象物体100を10°ごとに回転させ、36枚の物体画像A1～A36を得る。以下の説明では、このようにして得られた36枚の物体画像A1～A36をもとに立体モデル29を作成する場合を説明する。ここで、カメラ120の位置および俯角（または、仰角）は固定されている。また、カメラ120と回転テーブル110は、演算／制御部113によって制御される。なお、ステップS10で求めた背景画像および物体画像は、画像記憶部220に記憶される。本実施の形態12では、カメラを固定し、対象物体側を回転させて撮影しており、背景の撮影回数を可能な限り少なくするために背景を1回だけ撮影し、1枚の背景画像を得ている。しかし、より信

頼性のある背景画像を得るために、背景を2回以上撮影し、2枚以上の背景画像を得ても良い。

5       なお、本実施の形態12のように、カメラ120を固定し、対象物体100を回転させることによって、対象物体100の周りの複数方向から背景を含めて対象物体100を撮影する場合は背景の撮影は1回でもよいが、対象物体100を固定し、カメラ120を対象物体100を中心にして回すことによって、対象物体100の周りの複数方向から背景を含めて対象物体100を撮影する場合は背景の撮影は複数回必要である。

10       ステップS12では、図示しないシルエット作成部が、シルエット画像を作成する。つまり、各物体画像A1～A36と、背景画像との間で差分処理を行ない、複数のシルエット画像B1～Bnを作成する。物体画像A1～A36が36枚であるため、シルエット画像も36枚になる。ここで、差分処理（差分を求める処理）は、画素毎に、物体画像の色情報と、背景画像の色情報との差を求めることをいう。ステップS14では、図示しないボーディング部が、ボーディング処理  
15       を行なう。つまり、複数のシルエット画像B1～B36をもとに、円筒座標系ボクセル空間251へボーディング処理を行なう。そして、図示しないしきい値処理部（立体形状獲得部）は、投票数がしきい値以上の部分を、対象物体100の立体形状（存在領域）とする。

20       なお、ボクセル空間としては直交座標系ボクセル空間を用いてもよいが、円筒座標系ボクセル空間251を用いた方がメモリ量を小さく抑制しつつ良好な形状獲得が行なえる場合が多い。

25       ステップS16では、ステップS14で求めた対象物体100の立体形状をもとに、複数の立体形状表現要素（たとえば、三角パッチなどのポリゴン、以下、簡単のため立体形状表現要素をポリゴンと表記）27を作成し、ステップS14で求めた対象物体100の立体形状を複数のポリゴン27で表現する。ポリゴン27で表現された立体形状は、形状記憶部230に記憶される。ステップS18では、ステップS16で作成した、各ポリゴン27に対応したテクスチャを物体画像から獲得し、各ポリゴン27にマップする。また、テクスチャ（色情報）は、色彩情報記憶部240に記憶される。なお、ステップS12～S18の処理は、

演算／制御部 113 が行ない、シルエット作成部、ボーディング部、しきい値処理部は、演算／制御部 113 に含まれる。以下、ステップ S 8 におけるキャリブレーション、ステップ S 14 におけるボーディング処理およびステップ S 16 におけるポリゴン作成について詳細に説明する。

5

## (キャリブレーション)

キャリブレーションとして、カメラ 120 の内部パラメータ（パースペクティブ比）およびカメラ 120 と回転テーブル 110 の位置関係を求める。まず、カメラ 120 の内部パラメータ（パースペクティブ比）について説明する。図 34 は、カメラ 120 の内部パラメータ（パースペクティブ比）を説明するための図である。図 34 を参照して、基準ブロック 31 を、カメラ 120 によって撮影する。この場合に、基準ブロック 31 がスクリーン 33 にちょうど入るように撮影する。このときの、カメラ 120 と基準ブロック 31 との間の距離  $L$  を計測する。さらに、基準ブロック 31 の高さ  $T$  を計測しておく。パースペクティブ比は、基準ブロック 31 の高さ  $T$  を、距離  $L$  で割ったものである。つまり、パースペクティブ比は、 $T/L$  で表わされる。遠近法においては、画面に投影される物体サイズは、視点から物体までの距離に応じて拡大／縮小されるが、その拡大／縮小比率を決めるパラメータがパースペクティブ比である。

10

15

20

25

次に、カメラ 120 と回転テーブル 110 との位置関係の計測について説明する。図 35 A から図 35 C は、カメラと回転テーブルとの位置関係の計測を説明するための図である。図 35 A は、回転テーブル 110 の座標系（ $x y z$  座標系）に置かれたカメラ 120 を示す図である。図 35 A を参照して、回転テーブル 110 の座標系（ $x y z$  座標系）を用いて、カメラ 120 の位置（ $x_0, y_0, z_0$ ）を求める。さらに、カメラ 120 の光軸 35 の回りの回転角  $\alpha$  を求める。図 35 B は、図 35 A のカメラ 120 の  $y z$  平面への正射影を示す図である。図 35 B を参照して、カメラ 120 の光軸 35 と、 $y$  軸がなす角  $\beta$  を求める。図 35 C は、図 35 A のカメラ 120 の  $x y$  平面への正射影を示す図である。図 35 C を参照して、カメラ 120 の光軸 35 と  $y$  軸がなす角  $\gamma$  を求める。

つまり、カメラ 120 と回転テーブル 110 の位置関係として、回転テーブル 110 の座標系（ $x y z$  座標系）を用いたカメラ 120 の位置および角度  $\alpha, \beta,$

$\gamma$ を求める。なお、本実施の形態12では、角度 $\alpha$ 、 $\gamma$ は、ほぼ $0^\circ$ にしている。ここで、角度 $\beta$ は、回転テーブル110に対するカメラ120の俯角である。なお、角度 $\beta$ を、回転テーブル110に置かれた対象物体に対するカメラ120の俯角と呼ぶこともある。ここで、俯角には、負の俯角、つまり、仰角を含む。

- 5      以上のように、本実施の形態12では、キャリブレーションとして、対象物体に対するカメラの俯角を求めるため、この俯角をもって対象物体を撮影した物体画像をもとに立体モデル29を作成することができる。つまり、対象物体を横方向（x y 平面に平行な方向）から撮影して得られた物体画像だけでなく、斜め上から対象物体を撮影して得られた物体画像をもとに立体モデル29を作成できる。
- 10     このため、横方向のみからの撮影では得られないような対象物体の上面部も含め、色情報を十分に得ることができる。さらに、対象物体の局所的凹部をも認識できるので、精度よく立体モデル29を作成できる。

(ボーディング処理)

- 15     図33のステップS14におけるボーディング処理の詳細を説明する。図36は、ボーディング処理を施す円筒座標系ボクセル空間251を説明するための図である。図36を参照して、円筒座標系ボクセル空間251は、複数のボクセル39からなる。ここで、円筒座標系ボクセル空間251におけるボクセルを説明するための便宜のため、図36の円筒座標系ボクセル空間251を中心軸40を持つ円柱と考える。そして、この円柱25を、中心軸40に垂直な異なる複数の
- 20     平面で切断する。さらに、中心軸40を含み、かつ、中心軸40に平行な複数の平面で円柱25を切断する。さらに、中心軸40を軸とする異なる複数の回転面で円柱25を切断する。以上のようにして、円柱25を切断することによって得られた円柱25の各要素を考える。この各要素が、円筒座標系ボクセル空間251における各ボクセル39に対応する。

- 25     図37は、ボーディング処理を説明するための図である。図33のステップS12で得られた36枚のシルエット画像B1～B36をもとに、円筒座標系ボクセル空間251へボーディング処理を行なう。なお、図37では、2つのシルエット画像B1、B2だけを図示している。

ここで仮定存在領域50を考える。図7は、仮定存在領域を説明するための図

である。図 7 では、1 つのシルエット画像 B 1 だけを図示している。図 3 7 およ  
び図 7 を参照して、シルエット画像 B 1 に注目して、シルエット画像 B 1 に対す  
る仮定存在領域 5 0 とは、カメラの投影中心 5 1 を頂点とし、シルエット画像 B  
1 中の物体像 4 2 (対象物体 1 0 0 の輪郭) を断面形状とする錐体状の領域であ  
5 る。なお、他のシルエット画像 B 2 ~ B 3 6 に対する仮定存在領域も同様に定義  
できる。対象物体 1 0 0 は、この仮定存在領域の内側に必ず存在していることにな  
る。

図 3 6 も参照して、ボーティング処理では、仮定存在領域 5 0 に存在するボク  
セル 3 9 のすべてに「1」を投票する。このようなボーティング処理をすべての  
10 シルエット画像 B 1 ~ B 3 6 に対して行なう。たとえば、3 6 枚のシルエット画  
像 B 1 ~ B 3 6 に対応するすべての仮定存在領域が重なり合う部分に存在するボ  
クセル 3 9 における投票数は「3 6」になる。

図 3 3 のステップ S 1 0 では、1 0° ごとに対象物体を撮影し 3 6 枚の物体画  
像を得て、ステップ S 1 2 で 3 6 枚のシルエット画像 B 1 ~ B 3 6 を生成してい  
15 る。このため、仮定存在領域の頂点 (カメラの投影中心に相当) は、中心軸 4 0  
のまわりに、1 0° ごとに位置することになる。また、仮定存在領域の頂点 (カ  
メラの投影中心に相当) の位置は、図 3 3 のステップ S 8 におけるキャリブレー  
ションの結果に従って決定される。つまり、パースペクティブ比によって、シル  
エット画像 B 1 ~ B 3 6 とそれに対応する仮定存在領域の頂点 (カメラの投影中  
20 心に相当) との位置関係が決定される。すなわち、仮定存在領域である錐体の広  
がり角が決定される。また、カメラ 1 2 0 と回転テーブル 1 1 0 との位置関係に  
よって、シルエット画像 B 1 ~ B 3 6 に対応する仮定存在領域の頂点 (カメラの  
投影中心に相当) と円筒座標系ボクセル空間 2 5 1 との位置関係が決定される。

図 3 8 は、ボーティング処理の結果を示す図である。図 3 8 を参照して、色の  
25 濃い部分は投票数が多く、色の薄い部分は投票数が少ない。なお、図 3 8 の z 軸  
は、図 3 7 の中心軸 4 0 に相当する。

すべてのシルエット画像 B 1 ~ B 3 6 に対するボーティング処理が終了した後、  
しきい値処理を行なう。すなわち、所定のしきい値以上の投票数を有するボクセ  
ル 3 9 が存在する領域を対象物体 1 0 0 の存在領域とする。この存在領域の形状

が対象物体 100 の立体形状である。たとえば、しきい値を「32」とすると、投票数が「32」以上のボクセル 39 が存在する領域の形状が対象物体 100 の立体形状となる。

5 以上のように、本実施の形態 12 では、ボーディング処理によって、対象物体の立体形状を求めている。このため、ボーディング処理で用いる複数のシルエット画像のうちのいくつかが不正確であっても、精度よく立体モデル 29 を作成できる。なお、従来は複数の仮定存在領域の論理積により立体形状を得ているため、たとえば、シルエット画像中の物体像が不正確で、対象物体の輪郭を正確に表わしておらず、対象物体の形状の一部が欠けたようなものを表している場合は、  
10 その欠けた部分については、対象物体の立体形状として表現されないことになる。ここではボーディング処理でボクセル空間 251 における対象物体の存在領域を推定しているが、ボクセル空間 251 における対象物体の存在確率を求めることができれば、ボーディング処理以外のいかなる処理で対象物体の存在領域を推定するようにしてもよい。

15

#### (ポリゴン生成)

図 39 A および図 39 B は、図 33 のステップ S16 におけるポリゴン生成の具体的内容を説明するための図である。図 40 は、図 33 のステップ S16 におけるポリゴン生成の流れを説明するための図である。図 39 B は、図 39 A 中の 39 B 部に存在する輪郭線 43 a, 43 b に基づき求められたポリゴンを示す図  
20 である。図 39 A および図 40 を参照して、ステップ S A1 では、図示しない切断部が、ボーディング処理の結果に基づいて求められた対象物体 100 の立体形状 (図 38 参照) を複数の平面 (図 39 A では、3 つの平面 41 a, 41 b, 41 c だけを図示) で切断し、各切断面 (図 39 A では、3 つの切断面 44 a, 44 b, 44 c だけを図示) の輪郭線 (図 39 A では、3 つの輪郭線 43 a, 43  
25 b, 43 c だけを図示) を求める。ここで、図 33 のステップ S10 では 10° ごとに対象物体を撮影し物体画像を得て、ステップ S12 で 10° ごとのシルエット画像 B1 ~ B36 を生成している。このため、対象物体 100 の立体形状を中心軸 40 のまわりに 10° ごとに複数の平面で切断する。つまり、隣り合う平面同士の間す角度  $\theta$  が 10° になるような複数の平面で対象物体 100 の立体形状

を切断する。対象物体 1 0 0 の立体形状を切断する各平面は、中心軸 4 0 を含む平面である。

ステップ S A 2 では、図示しない多角形近似部は、各切断面の各輪郭線を多角形近似し、その多角形の頂点座標を求める。ここで、多角形近似の方法としては、  
5     たとえば、J. Ramer, "An Iterative Procedure for the Polygonal Approximation of Plane Curves", CGIP, Vol. 1, pp. 244-256, 1972 に開示されている方法などを用いることができる。そして、図示しない接続部は、各切断面において、隣り合う頂点同士を直線で接続する。ステップ S A 3 では、隣接する切断面間で、各切断面の輪郭線に対応する頂点同士を接続しポリ  
10     ゴンを生成する。なお、ステップ S A 2 の多角形近似においては、その近似精度を可変することにより、最終的に生成されるポリゴン数を制御することも可能である。

図 3 9 B を参照して、ステップ S A 2 およびステップ S A 3 における処理を説明する。ステップ S A 2 では、輪郭線 4 3 a, 4 3 b を多角形近似し、その多角  
15     形の頂点 4 5 a, 4 5 b の座標を求める。そして、輪郭線 4 3 a を多角形近似することによって求めた複数の頂点 4 5 a について、隣り合うもの同士を直線で接続する。輪郭線 4 3 b を多角形近似して求めた複数の頂点 4 5 b についても同様の処理を行なう。ここで、輪郭線 4 3 a に対応するのが頂点 4 5 a であり、輪郭線 4 3 b に対応するのが頂点 4 5 b である。ステップ S A 3 では、切断面 4  
20     4 a の輪郭線 4 3 a に対応する頂点 4 5 a と、切断面 4 4 b の輪郭線 4 3 b に対応する頂点 4 5 b との間を直線で接続しポリゴン 2 7 を生成する。ここで、頂点 4 5 a と頂点 4 5 b とを直線で接続する方法として、局所的最近接点接続戦略および大域的最近接点接続戦略がある。

局所的最近接点接続戦略は、隣接する切断面の一方の輪郭線を多角形近似することによって得られた頂点と、隣接する切断面の他方の輪郭線を多角形近似することによって得られた頂点との間で、頂点間の長さが最も短い頂点間を直線で接続するものである。大域的最近接点接続戦略は、隣接する切断面の一方の輪郭線を多角形近似することによって得られた頂点と、隣接する切断面の他方の輪郭線を多角形近似することによって得られた頂点とを、頂点間の長さの合計が最小になる  
25



ように直線で接続するものである。

局所的最近接点接続戦略について詳細に説明する。図41は、隣接する切断面の輪郭線に対応する頂点間の関係を示す図である。ここで、隣接する切断面として、切断面  $S_{cnt}$  および切断面  $S_{cnt+1}$  を考える。図41を参照して、頂点  $a, b, c, d, e, f$  は、切断面  $S_{cnt}$  の輪郭線を多角形近似して得られたものである。頂点  $A, B, C, D, E, F, G$  は、切断面  $S_{cnt+1}$  の輪郭線を多角形近似して得られたものである。なお、円筒座標系ボクセル空間  $251$  を利用したポリゴン生成を前提にしているため、頂点  $a$  と頂点  $A$  は同一の点であり、頂点  $f$  と頂点  $G$  は同一の点である。

図42は、局所的最近接点接続戦略を説明するための図である。図42を参照して、横方向は、切断面  $S_{cnt}$  の頂点  $a \sim f$  に対応付けられており、縦方向は、切断面  $S_{cnt+1}$  の頂点  $A \sim G$  に対応付けられている。そして、各格子点における数字（○の中に記載した数字）は、切断面  $S_{cnt}$  の輪郭線に対応する頂点  $a \sim f$ （図41）と、切断面  $S_{cnt+1}$  の輪郭線に対応する頂点  $A \sim G$ （図41）との間の距離である。たとえば、 $d$  と  $D$  の交点（ $d$  および  $D$ ）で決定される格子点）には、図41の頂点  $d$  と頂点  $D$  との間の距離が記されている。すなわち、図41の頂点  $d$  と頂点  $D$  との間の距離は、「2」である。

図41および図42を参照して、局所的最近接点接続戦略では、まず、初期ポリゴンを生成する。初期ポリゴン生成方法は以下の2つの方法が考えられる。初期ポリゴン生成のための第1の方法は、無条件に頂点  $b$   $B$  間を直線で接続するものである。初期ポリゴン生成の第2の方法は、頂点  $b$   $B$  間、頂点  $a$   $C$  間、頂点  $A$   $c$  間のうち、距離が最小のものを選択し、その頂点間を直線で接続するものである。図41および図42の例では、初期ポリゴン生成の2つの方法のうちいずれの方法においても、頂点  $b$   $B$  間が選択され、頂点  $b$   $B$  が直線で接続される。

次に、頂点  $c$   $B$  間または頂点  $b$   $C$  間のいずれを接続するかを考える。頂点  $b$   $C$  間の距離が、頂点  $c$   $B$  間の距離より短いので、頂点  $b$   $C$  間を直線で接続する。次に、頂点  $c$   $C$  間または頂点  $b$   $D$  間のいずれを接続するかを考える。頂点  $b$   $D$  間の距離と頂点  $c$   $C$  間の距離は等しいのでいずれを接続してもよいが、ここでは、頂点  $b$   $D$  間を直線で接続する。次に、頂点  $c$   $D$  間または頂点  $b$   $E$  間のいずれを接続

するかを考える。頂点  $c$   $D$ 間の距離が、頂点  $b$   $E$ 間の距離より短いので、頂点  $c$   $D$ 間を直線で接続する。以下、このような処理を繰返し、切断面  $S_{cnt}$  の輪郭線に対応する頂点と、切断面  $S_{cnt+1}$  の輪郭線に対応する頂点とを直線で接続していく。つまり、図 4 2 の各格子点において、右に位置する格子点に対応する頂点間の距離と、下に位置する格子点に対応する頂点間の距離とを比較し、短い距離が記されている格子点に対応する頂点間を直線で接続する。図 4 3 は、局所的最近接点接続戦略によって、図 4 1 の頂点  $a \sim f$  と頂点  $A \sim G$  とを接続して得られたポリゴンを示す図である。なお、図 4 1 と同様の部分については、同一の参照符号を付し、その説明を省略する。図 4 3 を参照して、局所的最近接点接続戦略に従って、頂点  $a \sim f$  と頂点  $A \sim G$  とが接続され、ポリゴン（三角パッチ）27 が形成される。

図 4 4 は、局所的最近接点接続戦略によるポリゴン生成の流れの一部を説明するための図である。図 4 5 は、局所的最近接点接続戦略によるポリゴン生成の流れの他の一部を説明するための図である。なお、ここでは、初期ポリゴン生成方法として、無条件に第 1 番目の頂点同士を接続する方法（初期ポリゴン生成の第 1 の方法）を用いた例を示している。図 4 4 を参照して、ステップ  $SB1$  では、変数  $S_{max}$  に、図 4 0 のステップ  $SA1$  で得られた切断面の数を代入する。さらに、変数  $S_{cnt}$  に、「0」を代入する。ステップ  $SB2$  では、変数  $V_{max}$  に、 $S_{cnt}$  番目の切断面における頂点数を代入する。さらに、変数  $V_{cnt}$  に「0」を代入する。ステップ  $SB3$  では、 $S_{cnt}$  番目の切断面における  $V_{cnt}$  番目の頂点と、 $S_{cnt}$  番目の切断面における  $V_{cnt+1}$  番目の頂点とを接続する。ここで、たとえば、図 4 1 に示すような頂点  $a \sim f$ 、 $A \sim G$  を考える場合、頂点  $a$ 、 $A$  は、第 0 番目の頂点であり、頂点  $b$ 、 $B$  は第 1 番目の頂点であり、頂点  $c$ 、 $C$  は第 3 番目の頂点である。ステップ  $SB4$  では、変数  $V_{cnt}$  に、 $V_{cnt+1}$  を代入する。ステップ  $SB5$  において、変数  $V_{cnt}$  が  $V_{max}-1$  以上の場合には、ステップ  $SB6$  に進む。一方、ステップ  $SB5$  において、変数  $V_{cnt}$  が、 $V_{max}-1$  より小さいときは、ステップ  $SB3$  に進む。ステップ  $SB6$  では、変数  $S_{cnt}$  に、 $S_{cnt+1}$  を代入する。ステップ  $SB7$  において、変数  $S_{cnt}$  が  $S_{max}$  以上のときは、図 4 5 のステップ  $SB8$  へ進む。一

方、ステップSB7において、変数 $S_{cnt}$ が $S_{max}$ より小さいときは、ステップSB2に進む。

図45を参照して、ステップSB8では、変数 $S_{cnt}$ に「0」を代入する。ステップSB9では、変数 $i_{max}$ に、 $S_{cnt}$ 番目の切断面における頂点数を代入する。さらに、変数 $j_{max}$ に、 $S_{cnt}+1$ 番目の切断面における頂点数を代入する。ステップSB10では、初期ポリゴンを生成する。ここでは、無条件に第1番目の頂点同士を接続する方法（初期ポリゴン生成の第1の方法）を用いた例を示しており、 $S_{cnt}$ 番目の切断面の1番目の頂点と、 $S_{cnt}+1$ 番目の切断面の1番目の頂点とを接続する。さらに、変数 $i$ に「1」を代入し、変数 $j$ に「1」を代入する。ステップSB11では、変数 $i\_n$ に $i+1$ を代入し、変数 $j\_n$ に $j+1$ を代入する。ステップSB12において、 $dist([S_{cnt}:i], [S_{cnt}+1:j\_n])$ は、 $S_{cnt}$ 番目の切断面の $i$ 番目の頂点と、 $S_{cnt}+1$ 番目の切断面の $j\_n$ 番目の頂点との間の距離を意味する。また、 $dist([S_{cnt}:i\_n], [S_{cnt}+1:j])$ は、 $S_{cnt}$ 番目の切断面の $i\_n$ 番目の頂点と、 $S_{cnt}+1$ 番目の切断面の $j$ 番目の頂点との間の距離を意味する。つまり、ステップSB12においては、 $S_{cnt}$ 番目の切断面の $i$ 番目の頂点と、 $S_{cnt}+1$ 番目の $j\_n$ 番目の頂点との間の距離が、 $S_{cnt}$ 番目の切断面の $i\_n$ 番目の頂点と、 $S_{cnt}+1$ 番目の切断面の $j$ 番目の頂点との間の距離以下の場合には、ステップSB13に進む。これ以外の場合には、ステップSB14に進む。

ステップSB13では、 $S_{cnt}$ 番目の切断面の $i$ 番目の頂点と、 $S_{cnt}+1$ 番目の切断面の $j\_n$ 番目の頂点とを接続する。さらに、変数 $j$ に $j\_n$ を代入する。ステップSB14では、 $S_{cnt}$ 番目の切断面の $i\_n$ 番目の頂点と、 $S_{cnt}+1$ 番目の切断面の $j$ 番目の頂点とを接続する。さらに、変数 $i$ に、 $i\_n$ を代入する。ステップSB15において、変数 $i$ が $i_{max}-1$ 以上の場合には、ステップSB17に進む。一方、変数 $i$ が $i_{max}-1$ より小さい場合には、ステップSB16に進む。ステップSB17では、 $S_{cnt}$ 番目の切断面の $i$ 番目の頂点と、 $S_{cnt}+1$ 番目の切断面の $j \sim j_{max}-1$ 番目の各頂点とを接続する。ステップSB16において、変数 $j$ が $j_{max}-1$ 以上の場合には、

ステップSB18に進む。一方、変数 $j$ が $j_{\max}-1$ より小さい場合には、ステップSB11に進む。ステップSB18では、 $Scnt+1$ 番目の切断面の $j$ 番目の頂点と、 $Scnt$ 番目の切断面の $i \sim i_{\max}-1$ 番目の各頂点とを接続する。ステップSB19では、変数 $Scnt$ に $Scnt+1$ を代入する。ステップSB20において、変数 $Scnt$ が $S_{\max}$ より小さい場合には、ステップSB9に進む。一方、変数 $Scnt$ が $S_{\max}$ 以上の場合には、処理を終了する。なお、切断面は0番から $S_{\max}-1$ 番まで存在するが、図45において $Scnt$ が $S_{\max}-1$ の場合、 $S_{\max}$ 番目の切断面の頂点を考慮しなければならない場合がある。この場合、 $S_{\max}$ 番目の切断面は0番目の切断面と同一とみなすことにする。

大域的最短接続戦略によるポリゴン生成について、図42を用いて詳細に説明する。格子点 $a \wedge$  ( $a$ と $A$ の交点)を始点とし、格子点 $f \wedge G$  ( $f$ と $G$ の交点)を終点とするような経路を考える。そして、格子点を経由するたびに、経由した格子点に割当てられた距離の値がペナルティとして加わるとし、ペナルティが最小となるような経路を求める。つまり、格子点 $a \wedge$ から格子点 $f \wedge G$ までの複数の経路のうち、最短となる経路を求める。このような最短となる経路は、総当り法、分枝限定法、Dijkstra アルゴリズム、 $A^*$  アルゴリズムなどを用いて求める。図42において、太い実線で示された経路が最小ペナルティ経路(最短経路)となる。そして、最小ペナルティ経路(最短経路)上に存在する格子点に対応する頂点間(図41参照)を接続する。たとえば、最小ペナルティ経路(太い実線)は、格子点 $b \wedge B$ 上を通っているので、図41の頂点 $b$ と $B$ とが接続される。図43は、図41の頂点 $a \sim f$ と頂点 $A \sim G$ とを、大域的最短接続戦略によって、直線で接続して得られたポリゴンを示している。

図46は、大域的最短接続戦略によるポリゴン生成の流れを説明するための図である。なお、図44および図45と同様のステップについては同一の参照符号を付しその説明は適宜省略する。図46を参照してステップSC9では、 $Scnt$ 番目の切断面の頂点と、 $Scnt+1$ 番目の切断面の頂点とを、接続距離が最短になるように接続する。ステップSC10では、変数 $Scnt$ に、 $Scnt+1$ を代入する。ステップSC11において、変数 $Scnt$ が $S_{\max}$ より小さい

場合には、ステップS C 9に進む。一方、変数S c n tがS m a x以上の場合には、処理を終了する。

5 以上のように、本実施の形態1 2では、円筒座標系ボクセル空間2 5 1を用いて対象物体の立体形状を求め、この立体形状を円筒座標系ボクセル空間2 5 1の中心軸に沿って、複数の平面で切断し、切断面の輪郭線をもとに形状モデル3 0 0を作成する。このため、直交座標系ボクセル空間を用いて形状モデル3 0 0を作成する場合に比べ、形状モデル作成のためのデータ量が少なくて済み、高速処理が可能になる。また、多角形近似および局所的最近接点接続戦略または大域的  
10 最短接続戦略を用いて形状モデル3 0 0を構成するポリゴンを生成する。このため、回転軸に垂直な複数の平面で、対象物体の立体形状を切断して形状モデル3 0 0を作成する従来技術に比べても、データ量が少なくなり、さらに高速処理が可能になる。つまり、実時間で形状モデル3 0 0を作成できる。さらに、多角形近似および局所的最近接点接続戦略または大域的  
15 最短接続戦略を用いて形状モデル3 0 0を構成するポリゴン2 7を生成する。このため、データ量が少なくなり、さらに高速処理が可能になる。

本実施の形態1 2では、これまでに述べてきた効果の他に、以下に示す効果をも奏する。すなわち、本実施の形態1 2では、3次元デジタイザを用いて形状モデルを作成する場合に比べ、人手による作業を少なくすることができる。さらに、  
20 本実施の形態1 2では、レーザを用いた計測を行っていない。このため、対象物体の材質の制限を受けず、暗室で計測が必要になるなどの撮影環境の制限が小さくなる。さらに、本実施の形態1 2では、回転テーブル1 1 0、カメラ1 2 0およびコンピュータ1 3 0という簡単な構成で、立体モデル2 9を作成できる。さらに、本実施の形態1 2では、差分処理を行なうことによりシルエット画像を生成し、これを用いて形状モデル3 0 0を作成する。このため、同一色の背景板  
25 などのような特殊な撮影環境が不要である。

以上、カメラ1台と回転テーブルを用い、1 0°おきに対象物体を撮影することで立体モデル2 9を作成する例について説明したが、カメラの台数、撮影視点の移動手段、撮影角度の刻み幅はこれに限定されるものではない。複数台のカメラを用いて対象物体を撮影し、立体モデル2 9を作成してもよい。これにより、

さらに高精度な形状獲得が可能となる。また、撮影視点の移動手段として、人手制御の回転台や、ロボットアームを用いてもよい。さらに、撮影角度の刻み幅は、複雑な物体ではより細かくしてもよいし、方向によって撮影角度の刻み角度を可変にしてもよい。すなわち、複雑な外形を持っている方向に関してはより細かい刻み幅で回転させ、撮影してもよい。また、対象物体の撮影角度の刻み幅を変更した場合は、ボクセル空間 251 を切断してポリゴン表現された形状モデル 300 に再構成するための切断面の変更も行なう。撮影角度と切断面は、関連して設定する。このようにすれば、撮影して得られたシルエット画像からの輪郭情報が、精度よくポリゴンデータに反映される。

図 47 は、図 3 のコンピュータ 130 に対象物体の立体モデル 29 を作成させるためのプログラム 301 を記録した CD-ROM を示す図である。図 47 を参照して、コンピュータ 130 は、CD-ROM 260 に記録されているプログラム 301 に従って、対象物体の立体モデル 29 を作成する。CD-ROM 260 に記録されているプログラム 301 は、対象物体のシルエット画像を生成するステップ S12 と、ボーテイング処理を行なうステップ S14 と、ポリゴン生成を行なうステップ S16 と、テクスチャマップを行なうステップ S18 とを含む。

本実施の形態 12 では、ボクセル表現された形状モデルからポリゴン表現された形状モデル 300 に再構成する場合に断面の多角形近似の技術を使用した。本発明はこれに限定されるわけではない。たとえば、ワイヤフレーム表現された形状モデル 300 に代えてメタボール表現された形状モデルを用いてもよい。また、本実施の形態 12 ではシルエット画像に基づいてボクセル空間 251 へのボーテイング処理を行なってポリゴンを生成したが、多角形近似法を用いてシルエット画像をポリゴンに変換するようにしてもよい。なお、この場合は、当然、シルエット画像は不正確であるので、手動操作による修正が必要となる。

## 請求の範囲

1. 形状モデルに対するテクスチャ情報付与方法であって、  
対象物体の形状を複数の立体形状構成要素の集合による形状モデルとして記述  
5 するステップと、

前記対象物体を異なる視点から撮影することにより獲得された複数の物体画像  
情報に基づいて、前記形状モデルに対するテクスチャ情報を、前記立体形状構成  
要素ごとに、前記各物体画像情報の前記立体形状構成要素に対するテクスチャ情  
報量に応じて付与するステップとを備える、テクスチャ情報付与方法。

- 10 2. 前記テクスチャ情報量は、

前記立体形状構成要素ごとに、前記各立体形状構成要素の面法線方向と、前記  
各物体画像情報を撮影した方向との一致の程度により表現される、請求項1記載  
のテクスチャ情報付与方法。

3. 前記テクスチャ情報量は、

- 15 前記立体形状構成要素ごとに、前記各物体画像情報に投影される前記立体形状  
構成要素の面積により表現される、請求項1記載のテクスチャ情報付与方法。

4. 形状モデルに対するテクスチャ情報付与方法であって、

対象物体の形状を複数の立体形状構成要素の集合による形状モデルとして記述  
するステップと、

- 20 前記対象物体を異なる視点から撮影することにより獲得された複数の物体画像  
情報に基づいて、前記形状モデルに対するテクスチャ情報を前記立体形状構成要  
素ごとに、前記各物体画像情報の前記立体形状構成要素に対するテクスチャ情報  
量と前記立体形状構成要素間のテクスチャ連続性との双方に応じて付与するステ  
ップとを備える、テクスチャ情報付与方法。

- 25 5. 前記テクスチャ情報を付与するステップにおいては、

前記テクスチャ情報量の増加に伴い減少し、かつ、前記立体形状構成要素間の  
テクスチャ連続性の向上に伴い減少する評価関数を極小化するように、前記立体  
形状構成要素ごとに対応付けられた前記物体画像情報から、前記形状モデルに対  
するテクスチャ情報を付与する、請求項4記載のテクスチャ情報付与方法。

6. 前記評価関数においては、

前記テクスチャ連続性は、注目する前記立体形状構成要素とそれに隣接する前記立体形状構成要素とにそれぞれ対応する前記物体画像情報の撮影位置および撮影方向の差の関数として表現される、請求項5記載のテクスチャ情報付与方法。

5 7. 前記評価関数においては、

前記テクスチャ連続性は、位置変化を伴って前記物体画像情報を撮影した場合に各物体画像情報に位置変化に対応したラベル番号を与えたとき、注目する前記立体形状構成要素に割当てられたラベル番号と、前記注目する立体形状構成要素に隣接する立体形状構成要素に割当てられたラベル番号との差の増加に伴い増加する関数として表現される、請求項5記載のテクスチャ情報付与方法。

10

8. 前記評価関数においては、

前記テクスチャ連続性は、規則的な位置変化を伴って前記物体画像情報を撮影した場合に各物体画像情報にラベル番号を与えたとき、注目する前記立体形状構成要素に割当てられたラベル番号と、前記注目する立体形状構成要素に隣接する立体形状構成要素に割当てられたラベル番号との差の増加に伴い増加する関数として表現される、請求項5記載のテクスチャ情報付与方法。

15

9. 前記評価関数においては、

前記テクスチャ情報量は、前記立体形状構成要素ごとに、前記各物体画像情報に投影される前記立体形状構成要素の面積の関数として表現される、請求項5記載のテクスチャ情報付与方法。

20

10. 前記評価関数においては、

前記テクスチャ情報量は、前記立体形状構成要素ごとに、前記各立体形状構成要素の面法線方向と、前記各物体画像情報を投影した方向との一致の程度の関数として表現される、請求項5記載のテクスチャ情報付与方法。

25

11. 前記評価関数は、

i 番目 (i : 自然数) の前記立体形状要素に割当てられたラベル番号と前記 i 番目の前記立体形状構成要素に隣接する立体形状構成要素に割当てられたラベル番号との差の全立体形状構成要素についての総和と、前記 i 番目の立体形状構成要素に割当てられたラベル番号に対応する前記物体画像情報に投影される、前記



i 番目の立体形状構成要素の面積の全立体形状構成要素についての総和との線形結合で表現される、請求項 5 記載のテクスチャ情報付与方法。

1 2. 形状モデルに対するテクスチャ情報付与方法であって、

5 対象物体の形状を複数の立体形状構成要素の集合による形状モデルとして記述するステップと、

位置変化を伴って複数の物体画像情報を投影した場合に前記各物体画像情報に位置変化に対応したラベル番号を与えたとき、前記各立体形状構成要素に対するテクスチャ情報量の増加に伴い減少し、かつ、前記各立体形状構成要素とそれに隣接する立体形状構成要素とにそれぞれ割当てられるテクスチャ情報のテクスチャ連続性の向上に伴い減少する評価関数を極小化するように、前記立体形状構成要素ごとに前記ラベル番号を対応付けるステップと、

10

前記対応付けられたラベル番号に対応する物体画像情報と、前記対応付けられたラベル番号を含む所定数のラベル番号に対応する物体画像情報とに基づいて、前記各物体画像情報に投影される前記立体形状構成要素の面積に応じた重み平均処理を行なうことにより、前記立体形状構成要素にテクスチャ情報を付与するステップとを備える、テクスチャ情報付与方法。

15

1 3. 形状モデルに対するテクスチャ情報付与方法であって、

対象物体の形状を複数の立体形状構成要素の集合による形状モデルとして記述するステップと、

20 規則的な位置変化を伴って複数の物体画像情報を投影した場合に前記各物体画像情報にラベル番号を与えたとき、前記各立体形状構成要素に対するテクスチャ情報量の増加に伴い減少し、かつ、前記各立体形状構成要素とそれに隣接する立体形状構成要素とにそれぞれ割当てられるテクスチャ情報のテクスチャ連続性の向上に伴い減少する評価関数を極小化するように、前記立体形状構成要素ごとに前記ラベル番号を対応付けるステップと、

25

前記対応付けられたラベル番号に対応する物体画像情報と、前記対応付けられたラベル番号を含む所定数のラベル番号に対応する物体画像情報とに基づいて、前記各物体画像情報に投影される前記立体形状構成要素の面積に応じた重み平均処理を行なうことにより、前記立体形状構成要素にテクスチャ情報を付与するス

ステップとを備える、テクスチャ情報付与方法。

14. 形状モデルに対するテクスチャ情報付与方法であって、

対象物体を異なる視点から撮影することにより複数の物体画像情報を獲得するステップと、

5 前記対象物体の形状を複数の立体形状構成要素の集合による形状モデルとして記述するステップと、

前記立体形状構成要素ごとに、前記複数の物体画像情報のそれぞれに投影された前記立体形状構成要素に対応する面積に応じた、複数の前記物体画像情報についての重み平均処理を行なうことにより得られたテクスチャ情報を付与するステップとを備える、テクスチャ情報付与方法。

10

15. コンピュータに形状モデルへのテクスチャ情報の付与をさせるためのプログラムを記録した媒体であって、

前記プログラムは、

15 対象物体の形状を複数の立体形状構成要素の集合による形状モデルとして記述するステップと、

前記対象物体を異なる視点から撮影することにより獲得された複数の物体画像情報に基づいて、前記形状モデルに対するテクスチャ情報を、前記立体形状構成要素ごとに、前記各物体画像情報の前記立体形状構成要素に対するテクスチャ情報量に応じて付与するステップとを含む、テクスチャ情報付与プログラムを記録した媒体。

20

16. 前記テクスチャ情報量は、

前記立体形状構成要素ごとに、前記各立体形状構成要素の面法線方向と、前記各物体画像情報を撮影した方向との一致の程度により表現される、請求項15記載のテクスチャ情報付与プログラムを記録した媒体。

25

17. 前記テクスチャ情報量は、

前記立体形状構成要素ごとに、前記各物体画像情報に投影される前記立体形状構成要素の面積により表現される、請求項15記載のテクスチャ情報付与プログラムを記録した媒体。

18. コンピュータに形状モデルへのテクスチャ情報の付与をさせるための

プログラムを記録した媒体であって、

前記プログラムは、

対象物体の形状を複数の立体形状構成要素の集合による形状モデルとして記述するステップと、

5 前記対象物体を異なる視点から撮影することにより獲得された複数の物体画像情報に基づいて、前記形状モデルに対するテクスチャ情報を前記立体形状構成要素ごとに、前記各物体画像情報の前記立体形状構成要素に対するテクスチャ情報量と前記立体形状構成要素間のテクスチャ連続性との双方に応じて付与するステップとを含む、テクスチャ情報付与プログラムを記録した媒体。

10 19. 前記テクスチャ情報を付与するステップにおいては、

前記テクスチャ情報量の増加に伴い減少し、かつ前記立体形状構成要素間のテクスチャ連続性の向上に伴い減少する評価関数を極小化するように、前記立体形状構成要素ごとに対応付けられた前記物体画像情報から、前記形状モデルに対するテクスチャ情報を付与する、請求項18記載のテクスチャ情報付与プログラムを記録した媒体。

15

20. 前記評価関数においては、

前記テクスチャ連続性は、注目する前記立体形状構成要素とそれに隣接する前記立体形状構成要素とにそれぞれ対応する前記物体画像情報の撮影位置および撮影方向の差の関数として表現される、請求項19記載のテクスチャ情報付与プログラムを記録した媒体。

20

21. 前記評価関数においては、

前記テクスチャ連続性は、位置変化を伴って前記物体画像情報を撮影した場合に各物体画像情報に位置変化に対応したラベル番号を与えたとき、注目する前記立体形状構成要素に割当てられたラベル番号と、前記注目する立体形状構成要素に隣接する立体形状構成要素に割当てられたラベル番号の差の増加に伴い増加する関数として表現される、請求項19記載のテクスチャ情報付与プログラムを記録した媒体。

25

22. 前記評価関数においては、

前記テクスチャ連続性は、規則的な位置変化を伴って前記物体画像情報を撮影

した場合に各物体画像情報にラベル番号を与えたとき、注目する前記立体形状構成要素に割当てられたラベル番号と、前記注目する立体形状構成要素に隣接する立体形状構成要素に割当てられたラベル番号の差の増加に伴い増加する関数として表現される、請求項 19 記載のテクスチャ情報付与プログラムを記録した媒体。

5        23.    前記評価関数においては、

テクスチャ情報量は、前記立体形状構成要素ごとに、前記各物体画像情報に投影される前記立体形状構成要素の面積の関数として表現される、請求項 19 記載のテクスチャ情報付与プログラムを記録した媒体。

24.    前記評価関数においては、

10        前記テクスチャ情報量は、

前記立体形状構成要素ごとに、前記各立体形状構成要素の面法線方向と、前記各物体画像情報を撮影した方向との一致の程度の関数として表現される、請求項 19 記載のテクスチャ情報付与プログラムを記録した媒体。

25.    前記評価関数は、

15         $i$  番目 ( $i$  : 自然数) の前記立体形状構成要素に割当てられたラベル番号と前記  $i$  番目の立体形状構成要素に隣接する立体形状構成要素に割当てられたラベル番号との差の全立体形状構成要素についての総和と、前記  $i$  番目の立体形状構成要素に割当てられたラベル番号に対応する前記物体画像情報に投影される、前記  $i$  番目の立体形状構成要素の面積の全立体形状構成要素についての総和との線形結合で表現される、請求項 19 記載のテクスチャ情報付与プログラムを記録した媒体。

26.    コンピュータに形状モデルへのテクスチャ情報の付与をさせるためのプログラムを記録した媒体であって、

前記プログラムは、

25        対象物体の形状を複数の立体形状構成要素の集合による形状モデルとして記述するステップと、

位置変化を伴って複数の物体画像情報を撮影した場合に前記各物体画像情報に位置変化に対応したラベル番号を与えたとき、前記各立体形状構成要素に対するテクスチャ情報量の増加に伴い減少し、かつ、前記各立体形状構成要素とそれに

隣接する立体形状構成要素とにそれぞれ割当てられるテクスチャ情報のテクスチャ連続性の向上に伴い減少する評価関数を極小化するように、前記立体形状構成要素ごとに前記ラベル番号を対応付けるステップと、

- 5 前記対応付けられたラベル番号に対応する物体画像情報と、前記対応付けられたラベル番号を含む所定数のラベル番号に対応する物体画像情報とに基づいて、前記各物体画像情報に投影される前記立体形状構成要素の面積に応じた重み平均処理を行なうことにより、前記立体形状構成要素にテクスチャ情報を付与するステップとを含む、テクスチャ情報付与プログラムを記録した媒体。

- 10 27. コンピュータに形状モデルへのテクスチャ情報の付与をさせるためのプログラムを記録した媒体であって、

前記プログラムは、

対象物体の形状を複数の立体形状構成要素の集合による形状モデルとして記述するステップと、

- 15 規則的な位置変化を伴って複数の物体画像情報を撮影した場合に前記各物体画像情報にラベル番号を与えたとき、前記各立体形状構成要素に対するテクスチャ情報量の増加に伴い減少し、かつ、前記各立体形状構成要素とそれに隣接する立体形状構成要素とにそれぞれ割当てられるテクスチャ情報のテクスチャ連続性の向上に伴い減少する評価関数を極小化するように、前記立体形状構成要素ごとに前記ラベル番号を対応付けるステップと、

- 20 前記対応付けられたラベル番号に対応する物体画像情報と、前記対応付けられたラベル番号を含む所定数のラベル番号に対応する物体画像情報とに基づいて、前記各物体画像情報に投影される前記立体形状構成要素の面積に応じた重み平均処理を行なうことにより、前記立体形状構成要素にテクスチャ情報を付与するステップとを含む、テクスチャ情報付与プログラムを記録した媒体。

- 25 28. コンピュータに形状モデルへのテクスチャ情報の付与をさせるためのプログラムを記録した媒体であって、

前記プログラムは、

対象物体を異なる視点から撮影することにより複数の物体画像情報を獲得するステップと、

前記対象物体の形状を複数の立体形状構成要素の集合による形状モデルとして記述するステップと、

- 前記立体形状構成要素ごとに、前記複数の物体画像情報のそれぞれに投影された前記立体形状構成要素に対応する面積に応じた、複数の前記物体画像情報についての重み平均処理を行なうことにより得られたテクスチャ情報を付与するステップとを含む、テクスチャ情報付与プログラムを記録した媒体。

29. 形状モデルに対するテクスチャ情報付与装置であって、

対象物体の形状を複数の立体形状構成要素の集合による形状モデルとして記述する手段と、

- 前記対象物体を異なる視点から撮影することにより獲得された複数の物体画像情報に基づいて、前記形状モデルに対するテクスチャ情報を、前記立体形状構成要素ごとに、前記各物体画像情報の前記立体形状構成要素に対するテクスチャ情報量に応じて付与する手段とを備える、テクスチャ情報付与装置。

30. 前記テクスチャ情報量は、

- 前記立体形状構成要素ごとに、前記各立体形状構成要素の面法線方向と、前記各物体画像情報を撮影した方向との一致の程度により表現される、請求項29記載のテクスチャ情報付与装置。

31. 前記テクスチャ情報量は、

- 前記立体形状構成要素ごとに、前記各物体画像情報に投影される前記立体形状構成要素の面積により表現される、請求項29記載のテクスチャ情報付与装置。

32. 形状モデルに対するテクスチャ情報付与装置であって、

対象物体の形状を複数の立体形状構成要素の集合による形状モデルとして記述する手段と、

- 前記対象物体を異なる視点から撮影することにより獲得された複数の物体画像情報に基づいて、前記形状モデルに対するテクスチャ情報を前記立体形状構成要素ごとに、前記各物体画像情報の前記立体形状構成要素に対するテクスチャ情報量と前記立体形状構成要素間のテクスチャ連続性との双方に応じて付与する手段とを備える、テクスチャ情報付与装置。

33. 前記テクスチャ情報を付与する手段においては、

前記テクスチャ情報量の増加に伴い減少し、かつ、前記立体形状構成要素間のテクスチャ連続性の向上に伴い減少する評価関数を極小化するように、前記立体形状構成要素ごとに対応付けられた前記物体画像情報から、前記形状モデルに対するテクスチャ情報を付与する、請求項 3 2 記載のテクスチャ情報付与装置。

5        3 4.    前記評価関数においては、

前記テクスチャ連続性は、注目する前記立体形状構成要素とそれに隣接する前記立体形状構成要素とにそれぞれ対応する前記物体画像情報の撮影位置および撮影方向の差の関数として表現される、請求項 3 3 記載のテクスチャ情報付与装置。

3 5.    前記評価関数においては、

10        前記テクスチャ連続性は、位置変化を伴って前記物体画像情報を撮影した場合に各物体画像情報に位置変化に対応したラベル番号を与えたとき、注目する前記立体形状構成要素に割当てられたラベル番号と、前記注目する立体形状構成要素に隣接する前記立体形状構成要素に割当てられたラベル番号との差の増加に伴い増加する関数として表現される、請求項 3 3 記載のテクスチャ情報付与装置。

15        3 6.    前記評価関数においては、

前記テクスチャ連続性は、規則的な位置変化を伴って前記物体画像情報を撮影した場合に各物体画像情報にラベル番号を与えたとき、注目する前記立体形状構成要素に割当てられたラベル番号と、前記注目する立体形状構成要素に隣接する前記立体形状構成要素に割当てられたラベル番号との差の増加に伴い増加する関数として表現される、請求項 3 3 記載のテクスチャ情報付与装置。

3 7.    前記評価関数においては、

前記テクスチャ情報量は、前記立体形状構成要素ごとに、前記各物体画像情報に投影される前記立体形状構成要素の面積の関数として表現される、請求項 3 3 記載のテクスチャ情報付与装置。

25        3 8.    前記評価関数においては、

前記テクスチャ情報量は、前記立体形状構成要素ごとに、前記各立体形状構成要素の面法線方向と、前記各物体画像情報を撮影した方向との一致の程度の関数として表現される、請求項 3 3 記載のテクスチャ情報付与装置。

3 9.    前記評価関数は、

- i 番目 (i : 自然数) の立体形状構成要素に割当てられたラベル番号と前記 i 番目の前記立体形状構成要素に隣接する立体形状構成要素に割当てられたラベル番号との差の全立体形状構成要素についての総和と、前記 i 番目の立体形状構成要素に割当てられたラベル番号に対応する前記物体画像情報に投影される、前記
- 5 i 番目の立体形状構成要素の面積の全立体形状構成要素についての総和との線形結合で表現される、請求項 33 記載のテクスチャ情報付与装置。

40. 形状モデルに対するテクスチャ情報付与装置であって、

対象物体の形状を複数の立体形状構成要素の集合による形状モデルとして記述する手段と、

- 10 位置変化を伴って複数の物体画像情報を撮影した場合に前記各物体画像情報に位置変化に対応したラベル番号を与えたとき、前記各立体形状構成要素に対するテクスチャ情報量の増加に伴い減少し、かつ、前記各立体形状構成要素とそれに隣接する立体形状構成要素とにそれぞれ割当てられるテクスチャ情報のテクスチャ連続性の向上に伴い減少する評価関数を極小化するように、前記立体形状構成
- 15 要素ごとに前記ラベル番号を対応付ける手段と、

- 前記対応付けられたラベル番号に対応する物体画像情報と、前記対応付けられたラベル番号を含む所定数のラベル番号に対応する物体画像情報とに基づいて、前記各物体画像情報に投影される前記立体形状構成要素の面積に応じた重み平均処理を行なうことにより、前記立体形状構成要素にテクスチャ情報を付与する手
- 20 段とを備える、テクスチャ情報付与装置。

41. 形状モデルに対するテクスチャ情報付与装置であって、

対象物体の形状を複数の立体形状構成要素の集合による形状モデルとして記述する手段と、

- 規則的な位置変化を伴って複数の物体画像情報を撮影した場合に前記各物体画像情報にラベル番号を与えたとき、前記各立体形状構成要素に対するテクスチャ情報量の増加に伴い減少し、かつ、前記各立体形状構成要素とそれに隣接する立体形状構成要素とにそれぞれ割当てられるテクスチャ情報のテクスチャ連続性の向上に伴い減少する評価関数を極小化するように、前記立体形状構成要素ごとに前記ラベル番号を対応付ける手段と、
- 25



前記対応付けられたラベル番号に対応する物体画像情報と、前記対応付けられたラベル番号を含む所定数のラベル番号に対応する物体画像情報とに基づいて、前記各物体画像情報に投影される前記立体形状構成要素の面積に応じた重み平均処理を行なうことにより、前記立体形状構成要素にテクスチャ情報を付与する手段とを備える、テクスチャ情報付与装置。

4 2. 形状モデルに対するテクスチャ情報付与装置であって、対象物体を異なる視点から撮影することにより複数の物体画像情報を獲得する手段と、

前記対象物体の形状を複数の立体形状構成要素の集合による形状モデルとして記述する手段と、

前記立体形状構成要素ごとに、前記複数の物体画像情報のそれぞれに投影された前記立体形状構成要素に対応する面積に応じた、複数の前記物体画像情報についての重み平均処理を行なうことにより得られたテクスチャ情報を付与する手段とを備える、テクスチャ情報付与装置。

4 3. 対象物体を撮影して得られた物体画像から、不所望部分を除くようにして物体部分を抽出する物体抽出装置であって、

前記物体画像を複数の領域に分割する領域分割手段と、

前記物体画像の各画素の情報を、前記領域ごとにまとめる処理を行なうことにより、前記物体画像中の物体部分を特定し、抽出する抽出手段とを備える、物体抽出装置。

4 4. 前記抽出手段において、前記物体画像の各画素の情報を、前記領域ごとにまとめる前記処理は、前記物体画像の各画素の情報を、前記領域ごとに平均化する処理である、請求項 4 3 に記載の物体抽出装置。

4 5. 前記抽出手段は、前記領域ごとにまとめられた前記各画素の情報に対し、しきい値処理を行なうことにより、前記物体画像中の物体部分を特定し、抽出する、請求項 4 3 または請求項 4 4 に記載の物体抽出装置。

4 6. 前記物体画像の前記各画素の情報は、前記対象物体の背景だけを撮影した背景画像と、前記物体画像との間で差分処理を行なって得られた差分情報である、請求項 4 3 から請求項 4 5 のいずれかに記載の物体抽出装置。

4 7. 前記抽出手段は、

前記対象物体の背景だけを撮影した背景画像と、前記物体画像との間で差分処理を行なう差分処理手段と、

5 前記各領域において、前記差分処理によって得られた差分の絶対値の平均値を求める平均値導出手段と、

前記領域における差分の絶対値の前記平均値と、予め定められた値とを比較し、前記平均値が前記予め定められた値以上の領域を物体部分として抽出するしきい値処理手段とを含む、請求項 4 3 から請求項 4 5 のいずれかに記載の物体抽出装置。

10 4 8. 前記抽出手段は、

前記物体画像の各領域内の画素の平均値を算出する平均値導出手段と、

前記物体画像の各領域内の画素の平均値とその領域に対応する前記背景画像の領域内の画素の平均値との間で差分処理を行なう差分処理手段と、

15 前記差分処理によって得られた差分の絶対値と、予め定められた値とを比較し、前記差分の絶対値が前記予め定められた値以上の領域を物体部分として抽出するしきい値処理手段とを含む、請求項 4 3 から請求項 4 5 のいずれかに記載の物体抽出装置。

4 9. 前記物体画像の各画素の情報は奥行情報である、請求項 4 3 から請求項 4 5 のいずれかに記載の物体抽出装置。

20 5 0. 対象物体を撮影して得られた物体画像から、不所望部分を除くようにして物体部分を抽出する物体抽出方法であって、

前記物体画像を複数の領域に分割する領域分割ステップと、

25 前記物体画像の各画素の情報を、前記領域ごとにまとめる処理を行なうことにより、前記物体画像中の物体部分を特定し、抽出する抽出ステップとを含む、物体抽出方法。

5 1. 前記抽出ステップにおいて、前記物体画像の各画素の情報を、前記領域ごとにまとめる前記処理は、前記物体画像の各画素の情報を、前記領域ごとに平均化する処理である、請求項 5 0 に記載の物体抽出方法。

5 2. 前記抽出ステップは、前記領域ごとにまとめられた前記各画素の情報

に対し、しきい値処理を行なうことにより、前記物体画像中の物体部分を特定し、抽出する、請求項 50 または請求項 51 に記載の物体抽出方法。

- 5 53. 前記物体画像の前記各画素の情報は、前記対象物体の背景だけを撮影した背景画像と、前記物体画像との間で差分処理を行なって得られた差分情報である、請求項 50 から請求項 52 のいずれかに記載の物体抽出方法。

54. 前記抽出ステップは、

前記対象物体の背景だけを撮影した背景画像と、前記物体画像との間で差分処理を行なう差分処理ステップと、

- 10 前記各領域において、前記差分処理によって得られた差分の絶対値の平均値を求める平均値導出ステップと、

前記領域における差分の絶対値の前記平均値と、予め定められた値とを比較し、前記平均値が前記予め定められた値以上の領域を物体部分として抽出するしきい値処理ステップとを含む、請求項 50 から請求項 52 のいずれかに記載の物体抽出方法。

- 15 55. 前記抽出ステップは、

前記物体画像の各領域内の画素の平均値を算出する平均値導出ステップと、

前記物体画像の各領域内の画素の平均値とその領域に対応する前記背景画像の領域内の画素の平均値との間で差分処理を行なう差分処理ステップと、

- 20 前記差分処理によって得られた差分の絶対値と、予め定められた値とを比較し、前記差分の絶対値が前記予め定められた値以上の領域を物体部分として抽出するしきい値処理ステップとを含む、請求項 50 から請求項 52 のいずれかに記載の物体抽出方法。

56. 前記物体画像の各画素の情報は奥行情報である、請求項 50 から請求項 52 のいずれかに記載の物体抽出方法。

- 25 57. コンピュータに、対象物体を撮影して得られた物体画像から、不所望部分を除くようにして物体部分を抽出させるためのプログラムを記録した媒体であって、

前記プログラムは、

前記物体画像を複数の領域に分割する領域分割ステップと、

前記物体画像の各画素の情報を、前記領域ごとにまとめる処理を行なうことにより、前記物体画像中の物体部分を特定し、抽出する抽出ステップとを含む、物体抽出プログラムを記録した媒体。

5 58. 前記抽出ステップにおいて、前記物体画像の各画素の情報を、前記領域ごとにまとめる前記処理は、前記物体画像の各画素の情報を、前記領域ごとに平均化する処理である、請求項57に記載の物体抽出プログラムを記録した媒体。

59. 前記抽出ステップは、前記領域ごとにまとめられた前記各画素の情報に対し、しきい値処理を行なうことにより、前記物体画像中の物体部分を特定し、抽出する、請求項57または請求項58に記載の物体抽出プログラムを記録した媒体。

60. 前記物体画像の前記各画素の情報は、前記対象物体の背景だけを撮影した背景画像と、前記物体画像との間で差分処理を行なって得られた差分情報である、請求項57から請求項59のいずれかに記載の物体抽出プログラムを記録した媒体。

15 61. 前記抽出ステップは、  
前記対象物体の背景だけを撮影した背景画像と、前記物体画像との間で差分処理を行なう差分処理ステップと、

前記各領域において、前記差分処理によって得られた差分の絶対値の平均値を求める平均値導出ステップと、

20 前記領域における差分の絶対値の前記平均値と、予め定められた値とを比較し、前記平均値が前記予め定められた値以上の領域を物体部分として抽出するしきい値処理ステップとを含む、請求項57から請求項59のいずれかに記載の物体抽出プログラムを記録した媒体。

62. 前記抽出ステップは、  
25 前記物体画像の各領域内の画素の平均値を算出する平均値導出ステップと、  
前記物体画像の各領域内の画素の平均値とその領域に対応する前記背景画像の領域内の画素の平均値との間で差分処理を行なう差分処理ステップと、

前記差分処理によって得られた差分の絶対値と、予め定められた値とを比較し、前記差分の絶対値が前記予め定められた値以上の領域を物体部分として抽出する

しきい値処理ステップとを含む、請求項 5 7 から請求項 5 9 のいずれかに記載の物体抽出プログラムを記録した媒体。

6 3. 前記物体画像の各画素の情報は奥行情報である、請求項 5 7 から請求項 5 9 のいずれかに記載の物体抽出プログラムを記録した媒体。

5 6 4. 対象物体を撮影して得られた物体画像から、不所望部分を除くようにして物体部分を抽出する物体抽出装置であって、

前記物体画像の奥行情報を算出する奥行情報算出手段と、

前記物体画像を複数の領域に分割する領域分割手段と、

前記奥行情報の平均値を各領域ごとに算出する平均値算出手段と、

10 前記複数の領域のうち前記平均値が予め定められた範囲内の領域を前記物体部分として抽出する抽出手段とを備える、物体抽出装置。

6 5. 対象物体を撮影して得られた物体画像から、不所望部分を除くようにして物体部分を抽出する物体抽出方法であって、

前記物体画像の奥行情報を算出するステップと、

15 前記物体画像を複数の領域に分割するステップと、

前記奥行情報の平均値を各領域ごとに算出するステップと、

前記複数の領域のうち前記平均値が予め定められた範囲内の領域を前記物体部分として抽出するステップとを含む、物体抽出方法。

20 6 6. コンピュータに、対象物体を撮影して得られた物体画像から、不所望部分を除くようにして物体部分を抽出させるためのプログラムを記録した媒体であって、

前記プログラムは、

前記物体画像の奥行情報を算出するステップと、

前記物体画像を複数の領域に分割するステップと、

25 前記奥行情報の平均値を各領域ごとに算出するステップと、

前記複数の領域のうち前記平均値が予め定められた範囲内の領域を前記物体部分として抽出するステップとを含む、物体抽出プログラムを記録した媒体。

6 7. 対象物体を撮影して得られた物体画像と、前記対象物体の背景だけを複数回撮影して得られた複数の背景画像とに基づいて、前記物体画像から背景部

- 分を除くようにして物体部分を抽出する物体抽出装置であって、  
前記物体画像と前記背景画像との差の絶対値を算出する差分手段と、  
前記物体画像のうち前記差の絶対値がしきい値よりも大きい部分を前記物体部分として抽出する抽出手段と、
- 5 前記複数の背景画像の画像情報分布に基づいて前記しきい値を統計的に決定するしきい値決定手段とを備える、物体抽出装置。
68. 対象物体を撮影して得られた物体画像と、前記対象物体の背景だけを複数回撮影して得られた複数の背景画像とに基づいて、前記物体画像から不所望部分を除くようにして物体部分を抽出する物体抽出装置であって、
- 10 前記複数の背景画像において互いに同じ座標に位置する画素の平均値および標準偏差を各画素ごとに算出する算出手段と、  
前記物体画像の各画素の値とその画素に対応する前記背景画像の画素の平均値との差の絶対値を算出する差分手段と、  
前記物体画像の画素のうち前記差の絶対値が前記標準偏差の所定倍よりも大きい画素を前記物体部分として抽出する抽出手段とを備える、物体抽出装置。
- 15 69. 対象物体を撮影して得られた物体画像と、前記対象物体の背景だけを複数回撮影して得られた複数の背景画像とに基づいて、前記物体画像から不所望部分を除くようにして物体部分を抽出する物体抽出装置であって、  
前記複数の背景画像において互いに同じ座標に位置する画素の平均値および標準偏差を各画素ごとに算出する平均／標準偏差算出手段と、  
20 前記物体画像を複数の領域に分割する領域分割手段と、  
前記物体画像の各領域内の各画素の値とその領域に対応する前記背景画像の領域内の対応する画素の平均値との差の絶対値を算出する差分手段と、  
前記差の絶対値の平均値を各領域ごとに算出する平均差算出手段と、  
25 前記標準偏差の平均値を各領域ごとに算出する平均標準偏差算出手段と、  
前記複数の領域のうち前記差の絶対値の平均値が前記標準偏差の平均値の所定倍よりも大きい領域を前記物体部分として抽出する抽出手段とを備える、物体抽出装置。
70. 対象物体を撮影して得られた物体画像と、前記対象物体の背景だけを

複数回撮影して得られた複数の背景画像とに基づいて、前記物体画像から不所望部分を除くようにして物体部分を抽出する物体抽出装置であって、

前記複数の背景画像において互いに同じ座標に位置する画素の平均値および標準偏差を各画素ごとに算出する平均／標準偏差算出手段と、

5 前記物体画像を複数の領域に分割する領域分割手段と、

前記物体画像の各領域内の画素の平均値を算出する平均算出手段と、

前記物体画像の各領域内の画素の平均値とその領域に対応する前記背景画像の領域内の画素の平均値との差の絶対値を算出する差分手段と、

前記差の絶対値の平均値を各領域ごとに算出する平均差算出手段と、

10 前記標準偏差の平均値を各領域ごとに算出する平均標準偏差算出手段と、

前記複数の領域のうち前記差の絶対値の平均値が前記標準偏差の平均値の所定倍よりも大きい領域を前記物体部分として抽出する抽出手段とを備える、物体抽出装置。

7.1. 対象物体を複数回撮影して得られた複数の物体画像と、前記対象物体  
15 の背景だけを複数回撮影して得られた複数の背景画像とに基づいて、前記物体画像から不所望部分を除くようにして物体部分を抽出する物体抽出装置であって、

前記複数の背景画像において互いに同じ座標に位置する画素の平均値および標準偏差を各画素ごとに算出する平均／標準偏差算出手段と、

20 前記複数の物体画像において互いに同じ座標に位置する画素の平均値を各画素ごとに算出する平均算出手段と、

前記物体画像を複数の領域に分割する領域分割手段と、

前記物体画像の各領域内の各画素の前記平均値とその領域に対応する前記背景画像の領域内の対応する画素の前記平均値との差の絶対値を算出する差分手段と、

前記差の絶対値の平均値を各領域ごとに算出する平均差算出手段と、

25 前記標準偏差の平均値を各領域ごとに算出する平均標準偏差算出手段と、

前記複数の領域のうち前記差の絶対値の平均値が前記標準偏差の平均値の所定倍よりも大きい領域を前記物体部分として抽出する抽出手段とを備える、物体抽出装置。

7.2. 対象物体を撮影して得られた物体画像と、前記対象物体の背景だけを

複数回撮影して得られた複数の背景画像とに基づいて、前記物体画像から不所望部分を除くようにして物体部分を抽出する物体抽出方法であって、

前記複数の背景画像の分布に基づいてしきい値を統計的に決定するステップと、  
前記物体画像と前記背景画像との差の絶対値を算出するステップと、

- 5 前記物体画像のうち前記差の絶対値が前記しきい値よりも大きい部分を前記物体部分として抽出するステップとを含む、物体抽出方法。

7 3. 対象物体を撮影して得られた物体画像と、前記対象物体の背景だけを複数回撮影して得られた複数の背景画像とに基づいて、前記物体画像から不所望部分を除くようにして物体部分を抽出する物体抽出方法であって、

- 10 前記複数の背景画像において互いに同じ座標に位置する画素の平均値および標準偏差を各画素ごとに算出するステップと、

前記物体画像の各画素の値とその画素に対応する前記背景画像の画素の平均値との差の絶対値を算出するステップと、

- 15 前記物体画像の画素のうち前記差の絶対値が前記標準偏差の所定倍よりも大きい画素を前記物体部分として抽出するステップとを含む、物体抽出方法。

7 4. 対象物体を撮影して得られた物体画像と、前記対象物体の背景だけを複数回撮影して得られた複数の背景画像とに基づいて、前記物体画像から背景部分を除くようにして物体部分を抽出する物体抽出方法であって、

- 20 前記複数の背景画像において互いに同じ座標に位置する画素の平均値および標準偏差を各画素ごとに算出するステップと、

前記物体画像を複数の領域に分割するステップと、

前記物体画像の各領域内の各画素の値とその領域に対応する前記背景画像の領域内の対応する画素の平均値との差の絶対値を算出するステップと、

- 25 前記差の絶対値の平均値を各領域ごとに算出するステップと、  
前記標準偏差の平均値を各領域ごとに算出するステップと、

前記複数の領域のうち前記差の絶対値の平均値が前記標準偏差の平均値の所定倍よりも大きい領域を前記物体部分として抽出するステップとを含む、物体抽出方法。

7 5. 対象物体を撮影して得られた物体画像と、前記対象物体の背景だけを



複数回撮影して得られた複数の背景画像とに基づいて、前記物体画像から不所望部分を除くようにして物体部分を抽出する物体抽出方法であって、

前記複数の背景画像において互いに同じ座標に位置する画素の平均値および標準偏差を各画素ごとに算出するステップと、

5 前記物体画像を複数の領域に分割するステップと、

前記物体画像の各領域内の画素の平均値を算出するステップと、

前記物体画像の各領域内の画素の平均値とその領域に対応する前記背景画像の領域内の画素の平均値との差の絶対値を算出するステップと、

前記差の絶対値の平均値を各領域ごとに算出するステップと、

10 前記標準偏差の平均値を各領域ごとに算出するステップと、

前記複数の領域のうち前記差の絶対値の平均値が前記標準偏差の平均値の所定倍よりも大きい領域を前記物体部分として抽出するステップとを含む、物体抽出方法。

7 6. 対象物体を複数回撮影して得られた複数の物体画像と、前記対象物体の背景だけを複数回撮影して得られた複数の背景画像とに基づいて、前記物体画像から不所望部分を除くようにして物体部分を抽出する物体抽出方法であって、

前記複数の背景画像において互いに同じ座標に位置する画素の平均値および標準偏差を各画素ごとに算出するステップと、

20 前記複数の物体画像において互いに同じ座標に位置する画素の平均値を各画素ごとに算出するステップと、

前記物体画像を複数の領域に分割するステップと、

前記物体画像の各領域内の各画素の前記平均値とその領域に対応する前記背景画像の領域内の対応する画素の前記平均値との差の絶対値を算出するステップと、

前記差の絶対値の平均値を各領域ごとに算出するステップと、

25 前記標準偏差の平均値を各領域ごとに算出するステップと、

前記複数の領域のうち前記差の絶対値の平均値が前記標準偏差の平均値の所定倍よりも大きい領域を前記物体部分として抽出するステップとを含む、物体抽出方法。

7 7. コンピュータに、対象物体を撮影して得られた物体画像と、前記対象

物体の背景だけを複数回撮影して得られた複数の背景画像とに基づいて、前記物体画像から不所望部分を除くようにして物体部分を抽出させるためのプログラムを記録した媒体であって、

前記プログラムは、

- 5      前記複数の背景画像の分布に基づいてしきい値を統計的に決定するステップと、  
前記物体画像と前記背景画像との差の絶対値を算出するステップと、

前記物体画像のうち前記差の絶対値が前記しきい値よりも大きい部分を前記物体部分として抽出するステップとを含む、物体抽出プログラムを記録した媒体。

78.      コンピュータに、対象物体を撮影して得られた物体画像と、前記対象  
10      物体の背景だけを複数回撮影して得られた複数の背景画像とに基づいて、前記物体画像から不所望部分を除くようにして物体部分を抽出させるためのプログラムを記録した媒体であって、

前記プログラムは、

- 前記複数の背景画像において互いに同じ座標に位置する画素の平均値および標準偏差を各画素ごとに算出するステップと、  
15

前記物体画像の各画素の値とその画素に対応する前記背景画像の画素の平均値との差の絶対値を算出するステップと、

- 前記物体画像の画素のうち前記差の絶対値が前記標準偏差の所定倍よりも大きい画素を前記物体部分として抽出するステップとを含む、物体抽出プログラムを  
20      記録した媒体。

79.      コンピュータに、対象物体を撮影して得られた物体画像と、前記対象物体の背景だけを複数回撮影して得られた複数の背景画像とに基づいて、前記物体画像から不所望部分を除くようにして物体部分を抽出させるためのプログラムを記録した媒体であって、

- 25      前記プログラムは、

前記複数の背景画像において互いに同じ座標に位置する画素の平均値および標準偏差を各画素ごとに算出するステップと、

前記物体画像を複数の領域に分割するステップと、

前記物体画像の各領域内の各画素の値とその領域に対応する前記背景画像の領

域内の対応する画素の平均値との差の絶対値を算出するステップと、

前記差の絶対値の平均値を各領域ごとに算出するステップと、

前記標準偏差の平均値を各領域ごとに算出するステップと、

- 5 前記複数の領域のうち前記差の絶対値の平均値が前記標準偏差の平均値の所定倍よりも大きい領域を前記物体部分として抽出するステップとを含む、物体抽出プログラムを記録した媒体。

80. コンピュータに、対象物体を撮影して得られた物体画像と、前記対象物体の背景だけを複数回撮影して得られた複数の背景画像とに基づいて、前記物体画像から不所望部分を除くようにして物体部分を抽出させるためのプログラムを記録した媒体であって、
- 10

前記プログラムは、

前記複数の背景画像において互いに同じ座標に位置する画素の平均値および標準偏差を各画素ごとに算出するステップと、

前記物体画像を複数の領域に分割するステップと、

- 15 前記物体画像の各領域内の画素の平均値を算出するステップと、

前記物体画像の各領域内の画素の平均値とその領域に対応する前記背景画像の領域内の画素の平均値との差の絶対値を算出するステップと、

前記差の絶対値の平均値を各領域ごとに算出するステップと、

前記標準偏差の平均値を各領域ごとに算出するステップと、

- 20 前記複数の領域のうち前記差の絶対値の平均値が前記標準偏差の平均値の所定倍よりも大きい領域を前記物体部分として抽出するステップとを含む、物体抽出プログラムを記録した媒体。

81. コンピュータに、対象物体を複数回撮影して得られた複数の物体画像と、前記対象物体の背景だけを複数回撮影して得られた複数の背景画像とに基づいて、前記物体画像から不所望部分を除くようにして物体部分を抽出させるためのプログラムを記録した媒体であって、
- 25

前記プログラムは、

前記複数の背景画像において互いに同じ座標に位置する画素の平均値および標準偏差を各画素ごとに算出するステップと、

前記複数の物体画像において互いに同じ座標に位置する画素の平均値を各画素ごとに算出するステップと、

前記物体画像を複数の領域に分割するステップと、

5 前記物体画像の各領域内の各画素の前記平均値とその領域に対応する前記背景画像の領域内の対応する画素の前記平均値との差の絶対値を算出するステップと、  
前記差の絶対値の平均値を各領域ごとに算出するステップと、

前記標準偏差の平均値を各領域ごとに算出するステップと、

前記複数の領域のうち前記差の絶対値の平均値が前記標準偏差の平均値の所定倍よりも大きい領域を前記物体部分として抽出するステップとを含む、物体抽出  
10 プログラムを記録した媒体。

8 2. 対象物体の立体モデルを作成する立体モデル作成装置であって、

前記対象物体の背景を撮影し、かつ、前記背景を含めて前記対象物体を撮影する撮影手段と、

15 前記背景だけを撮影して得られた背景画像と、前記背景を含めて前記対象物体を撮影して得られた複数の物体画像との差分を求め、複数のシルエット画像を生成するシルエット生成手段と、

前記複数のシルエット画像を用いて、前記対象物体の立体モデルを作成する手段とを備える、立体モデル作成装置。

20 8 3. 前記対象物体を回転させるための回転手段をさらに備える、請求項 8 2 に記載の立体モデル作成装置。

8 4. 対象物体の立体モデルを作成する立体モデル作成装置であって、

前記対象物体の複数のシルエット画像を生成するシルエット生成手段と、

前記複数のシルエット画像をもとにボクセル空間における前記対象物体の存在領域を推定する推定手段と、

25 前記推定手段によって得られた前記対象物体の存在領域を用いて、前記対象物体の立体モデルを作成する手段とを備える、立体モデル作成装置。

8 5. 前記推定手段は前記ボクセル空間へボーディング処理を行なう、請求項 8 4 に記載の立体モデル作成装置。

8 6. 前記ボーディング処理の結果、投票数が所定のしきい値以上の部分を、

前記対象物体の前記存在領域とするしきい値処理手段をさらに備えた、請求項 8  
5 に記載の立体モデル作成装置。

87. 対象物体の立体モデルを作成する立体モデル作成方法であって、  
撮影装置によって前記対象物体の背景だけを撮影し、背景画像を得るステップ  
5 と、

前記対象物体を前記撮影装置によって前記背景を含めて撮影し、複数の物体画  
像を得るステップと、

前記背景画像と、前記複数の物体画像との差分を求め、複数のシルエット画像  
を生成するステップと、

10 前記複数のシルエット画像を用いて、前記対象物体の立体モデルを作成するス  
テップとを含む、立体モデル作成方法。

88. 前記対象物体を回転させるステップをさらに含む、請求項 17 または  
請求項 87 に記載の立体モデル作成方法。

89. 対象物体の立体モデルを作成する立体モデル作成方法であって、  
15 前記対象物体の複数のシルエット画像を生成するステップと、  
前記複数のシルエット画像をもとにボクセル空間における前記対象物体の存在  
領域を推定するステップと、

前記推定された前記対象物体の存在領域を用いて、前記立体モデルを作成する  
ステップとを含む、立体モデル作成方法。

20 90. 前記推定するステップは、前記ボクセル空間へボーテイング処理を行  
なう、請求項 89 に記載の立体モデル作成方法。

91. 前記ボーテイング処理の結果、投票数が所定のしきい値以上の部分を、  
前記対象物体の前記存在領域とするステップをさらに含む、請求項 90 に記載の  
立体モデル作成方法。

25 92. コンピュータに対象物体の立体モデルを作成させるためのプログラム  
を記録した媒体であって、

前記プログラムは、

前記対象物体の複数のシルエット画像を生成するステップと、

前記複数のシルエット画像をもとにボクセル空間における前記対象物体の存在

領域を推定するステップと、

前記複数のシルエット画像をもとにボクセル空間における前記対象物体の存在領域を推定するステップと、

5 前記推定された前記対象物体の存在領域を用いて、前記立体モデルを作成するステップとを含む、立体モデル作成プログラムを記録した媒体。

9 3. 前記プログラム中の前記推定するステップは、前記ボクセル空間へボーディング処理を行なう、請求項 9 2 に記載の立体モデル作成プログラムを記録した媒体。

9 4. 前記プログラムは、

10 前記ボーディング処理の結果、投票数が所定のしきい値以上の部分を、前記対象物体の前記存在領域とするステップをさらに含む、請求項 9 3 に記載の立体モデル作成プログラムを記録した媒体。

9 5. 対象物体を撮影して得られた物体画像と、前記対象物体の背景だけを複数回撮影して得られた複数の背景画像とに基づいて、前記物体画像から不所望部分を除くようにして物体部分を抽出する物体抽出装置であって、

15 前記複数の背景画像において互いに同じ座標に位置する画素の平均値および標準偏差を各画素ごとに算出する平均／標準偏差算出手段と、

前記物体画像を複数の領域に分割する領域分割手段と、

20 前記物体画像の各領域内の画素の平均値を算出するとともに、背景画像の画素の平均値の各領域内における平均値を算出する平均算出手段と、

前記物体画像の各領域内の画素の平均値とその領域に対応する前記背景画像の領域内の画素の領域内における平均値との差の絶対値を算出する差分手段と、

前記標準偏差の平均値を各領域ごとに算出する平均標準偏差算出手段と、

25 前記複数の領域のうち差の絶対値が標準偏差の平均値の所定倍よりも大きい領域を物体部分として抽出する抽出手段とを備える、物体抽出装置。

9 6. 対象物体を撮影して得られた物体画像と、前記対象物体の背景だけを複数回撮影して得られた複数の背景画像とに基づいて、前記物体画像から不所望部分を除くようにして物体部分を抽出する物体抽出方法であって、

前記複数の背景画像において互いに同じ座標に位置する画素の平均値および標

標準偏差を各画素ごとに算出するステップと、

前記物体画像を複数の領域に分割するステップと、

前記物体画像の各領域内の画素の平均値を算出するとともに、背景画像の画素の平均値の各領域内における平均値を算出するステップと、

- 5 前記物体画像の各領域内の画素の平均値とその領域に対応する前記背景画像の領域内の画素の領域内における平均値との差の絶対値を算出するステップと、

前記標準偏差の平均値を各領域ごとに算出するステップと、

前記複数の領域のうち差の絶対値が標準偏差の平均値の所定倍よりも大きい領域を物体部分として抽出するステップとを含む、物体抽出方法。

- 10 97. コンピュータに、対象物体を撮影して得られた物体画像と、前記対象物体の背景だけを複数回撮影して得られた複数の背景画像とに基づいて、前記物体画像から不所望部分を除くようにして物体部分を抽出させるためのプログラムを記録した媒体であって、

- 15 前記複数の背景画像において互いに同じ座標に位置する画素の平均値および標準偏差を各画素ごとに算出するステップと、

前記物体画像を複数の領域に分割するステップと、

前記物体画像の各領域内の画素の平均値を算出するとともに、背景画像の画素の平均値の各領域内における平均値を算出するステップと、

- 20 前記物体画像の各領域内の画素の平均値とその領域に対応する前記背景画像の領域内の画素の領域内における平均値との差の絶対値を算出するステップと、

前記標準偏差の平均値を各領域ごとに算出するステップと、

前記複数の領域のうち差の絶対値が標準偏差の平均値の所定倍よりも大きい領域を物体部分として抽出するステップとを含む、物体抽出プログラムを記録した媒体。

FIG. 1

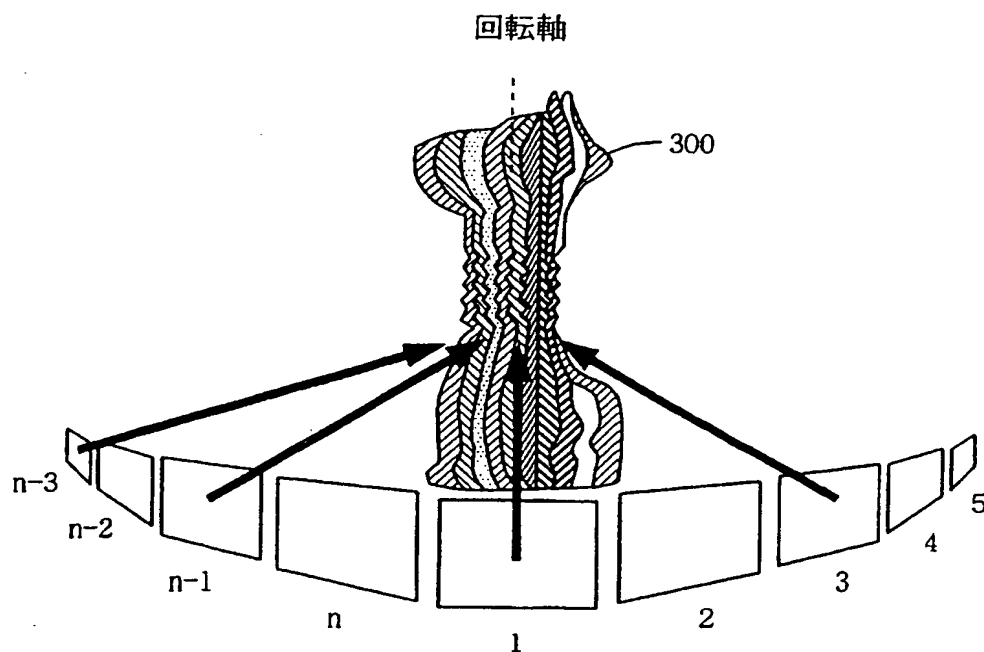


FIG. 2

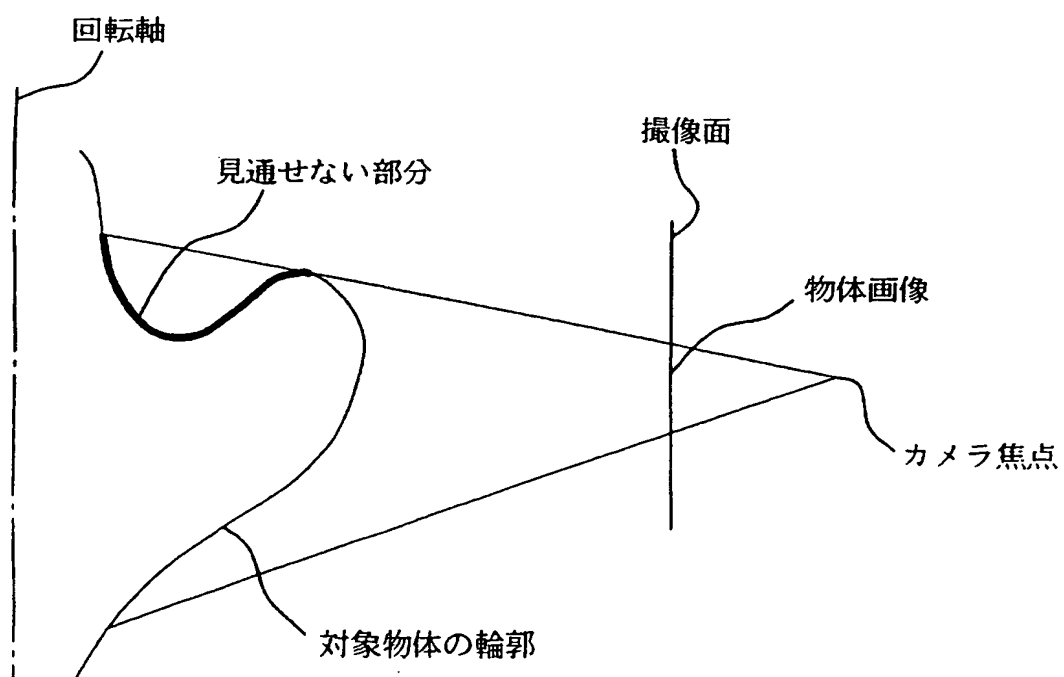




FIG. 3

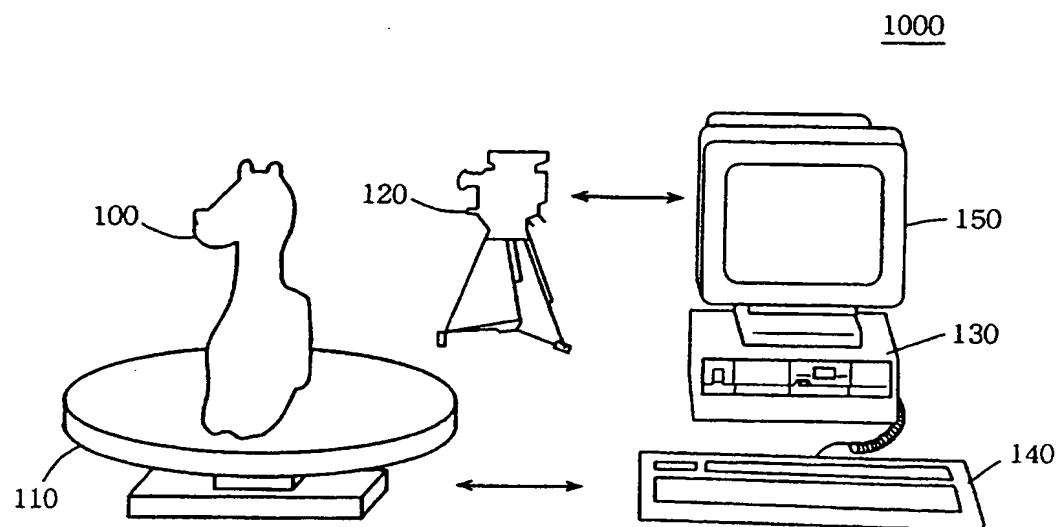


FIG. 4

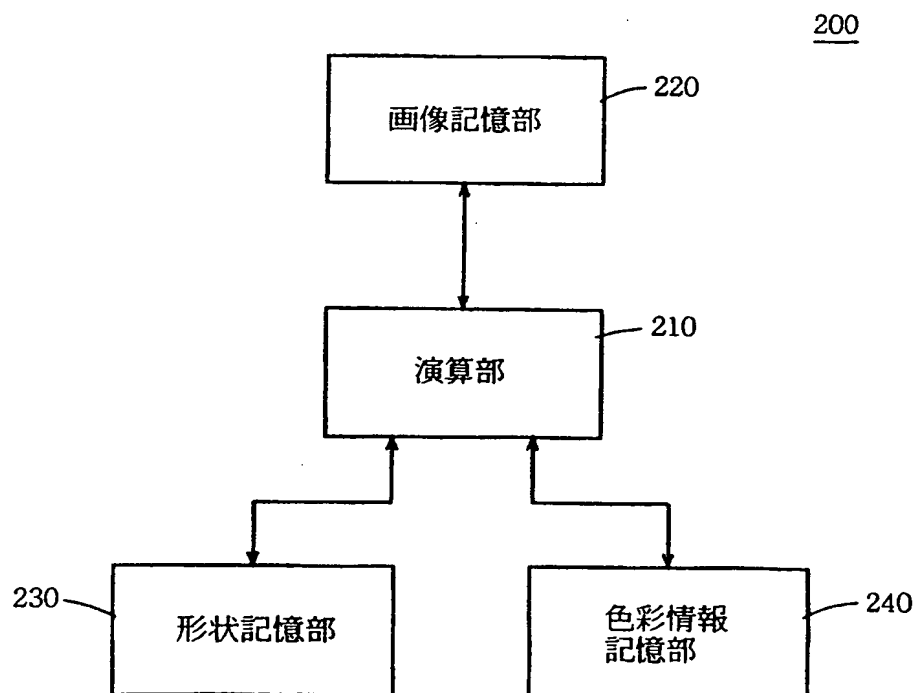


FIG. 5

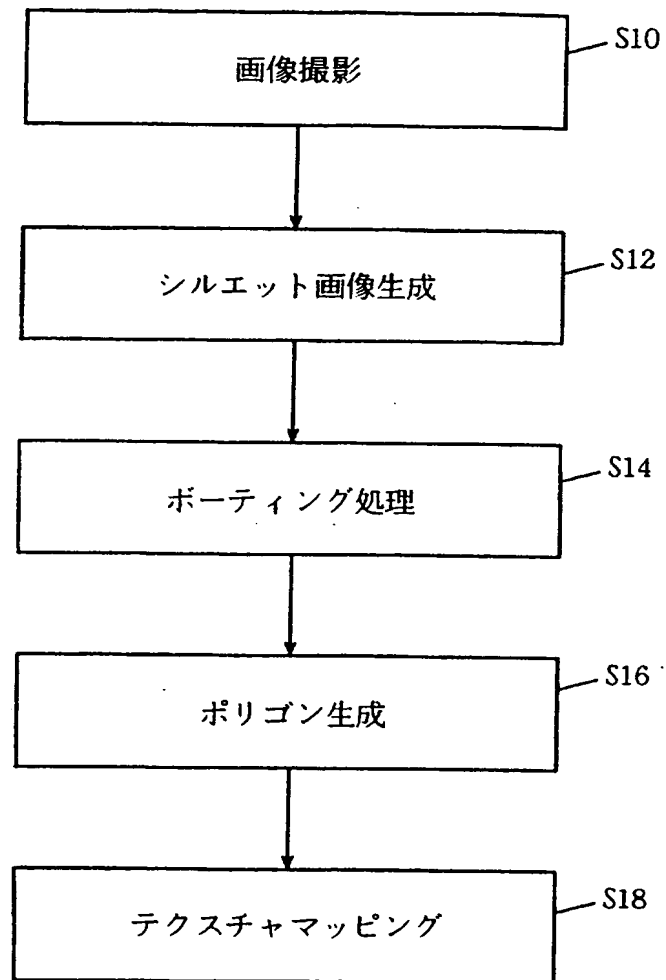


FIG. 6A

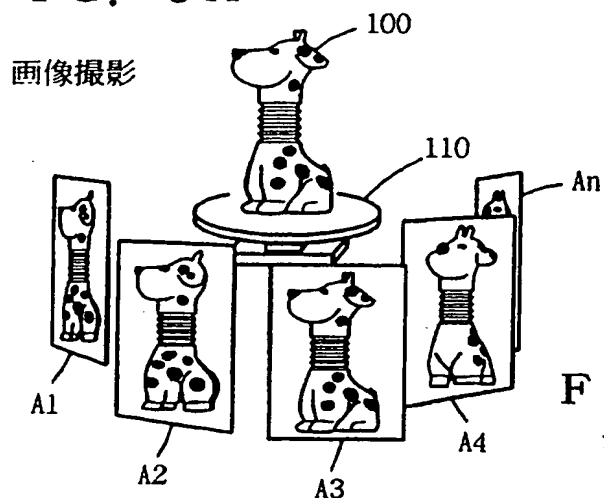


FIG. 6B

シルエット画像生成

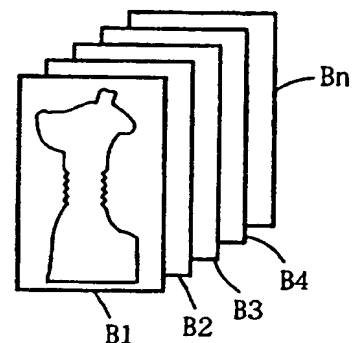


FIG. 6C

ボーディング処理

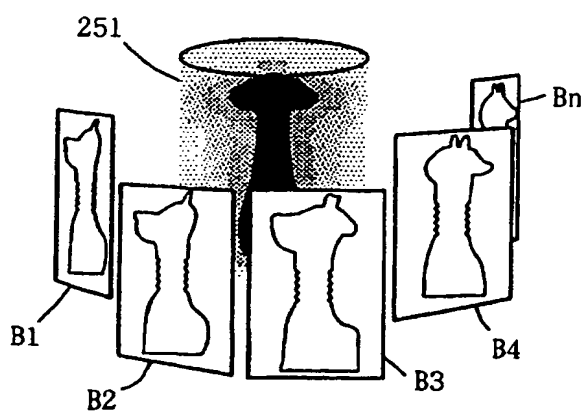


FIG. 6D

ポリゴン生成

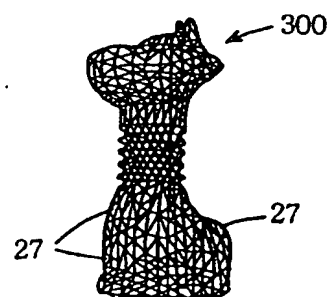


FIG. 6E

テクスチャマッピング

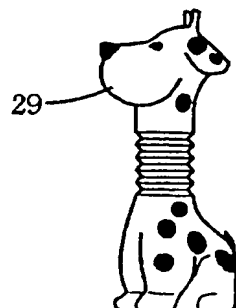


FIG. 7

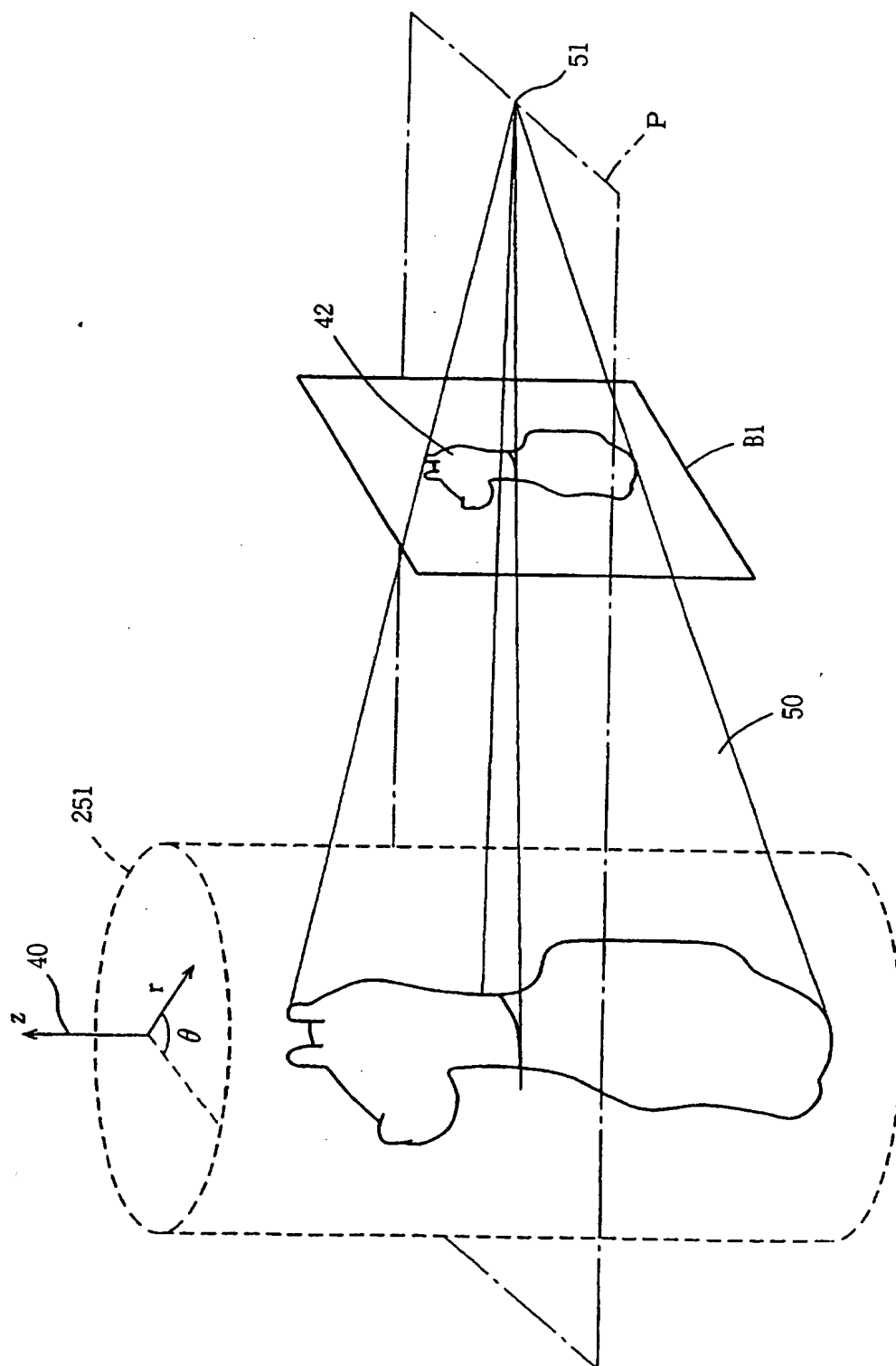


FIG. 8

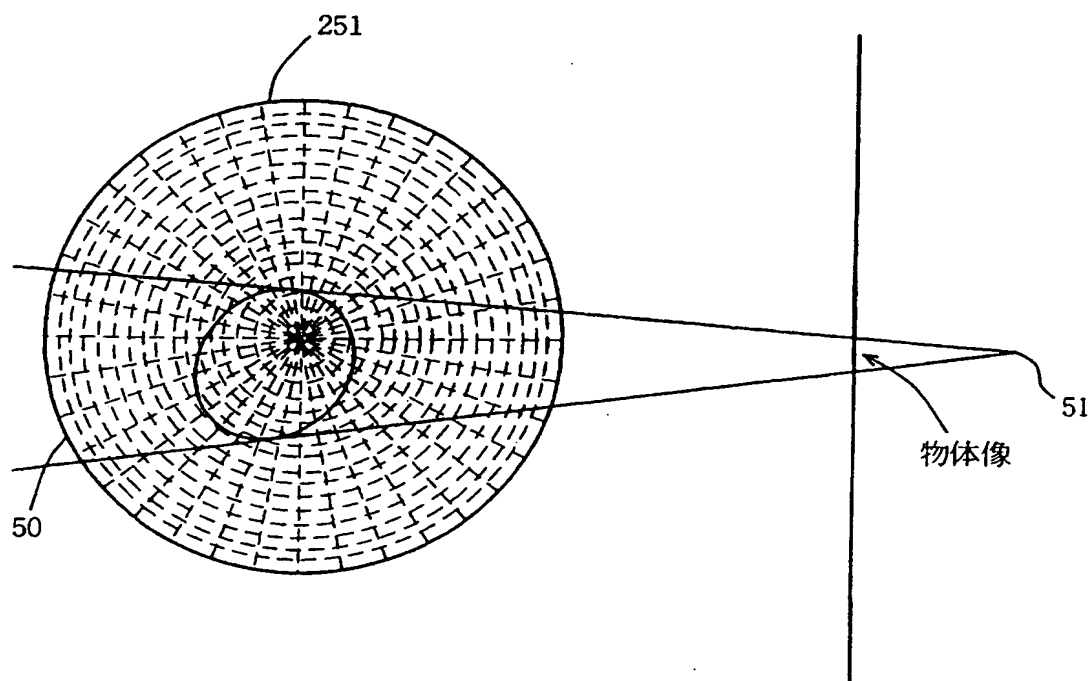


FIG. 9

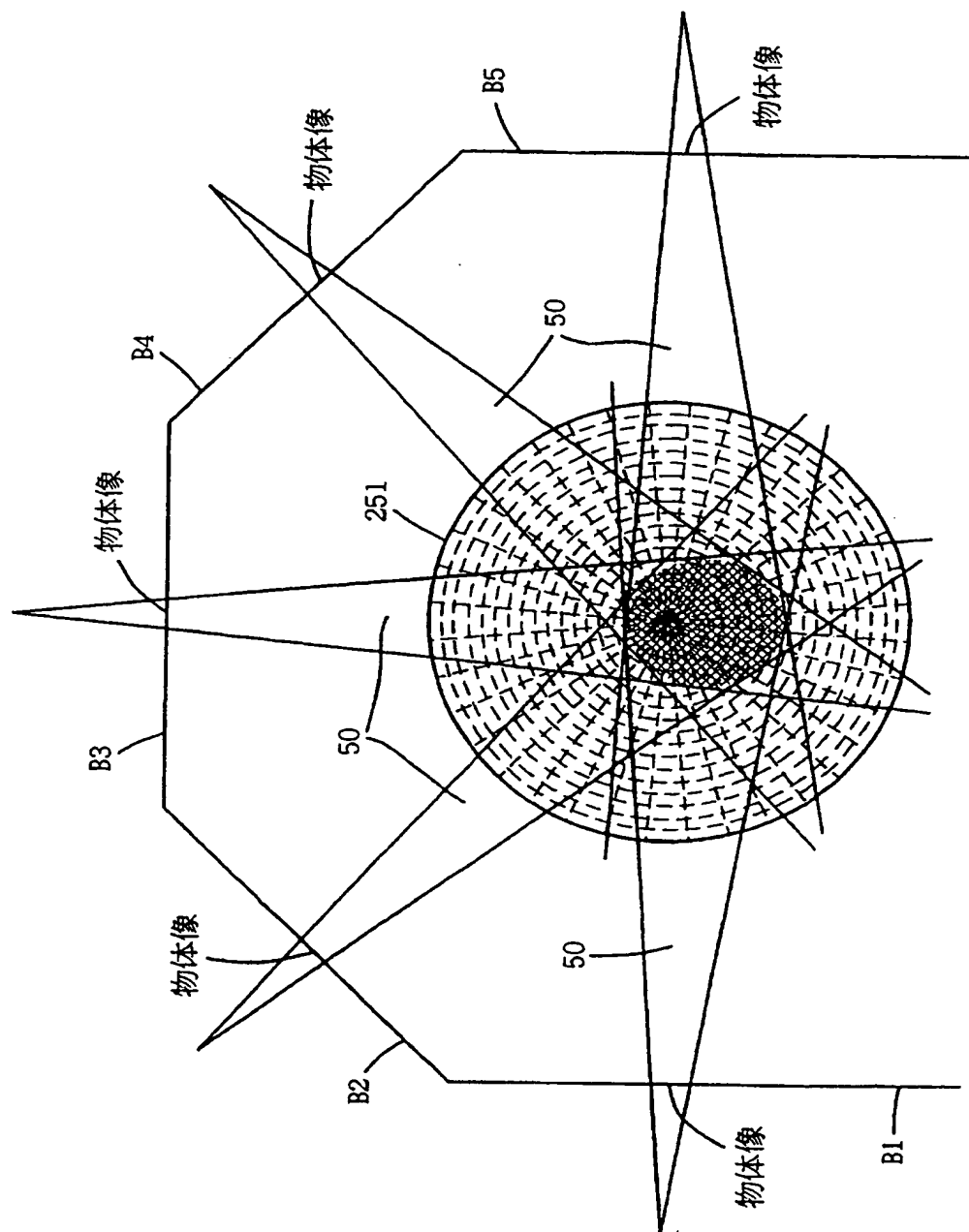


FIG. 10A

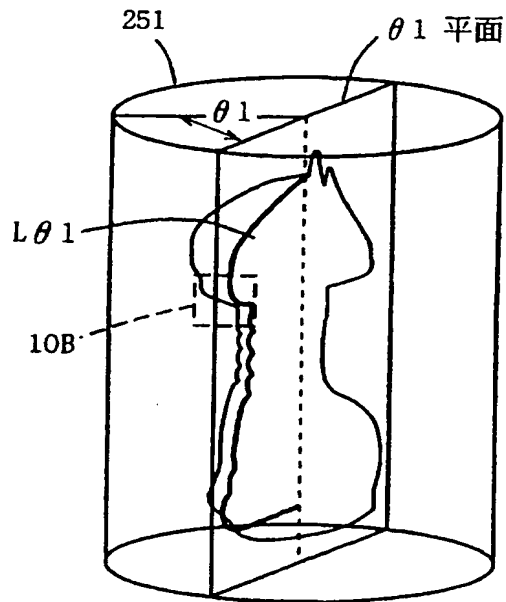


FIG. 10B

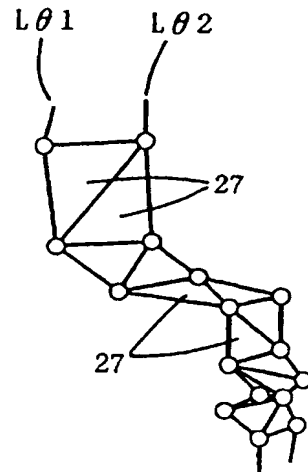


FIG. 11

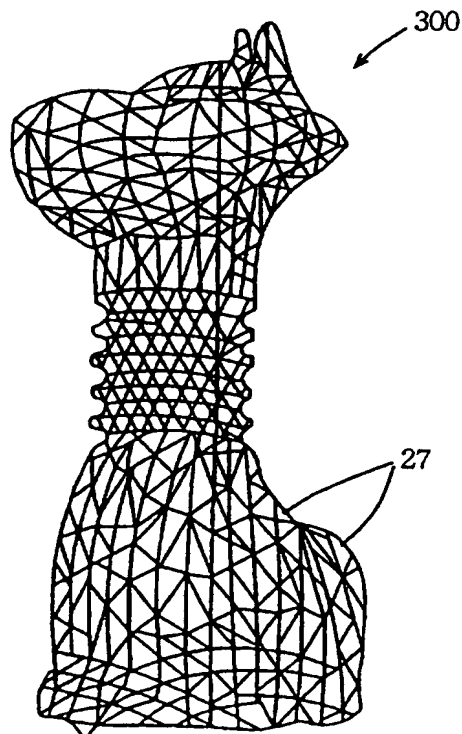


FIG. 12B

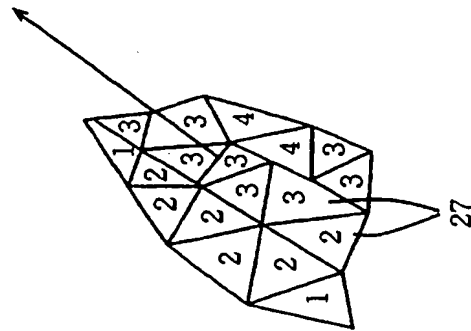


FIG. 12A

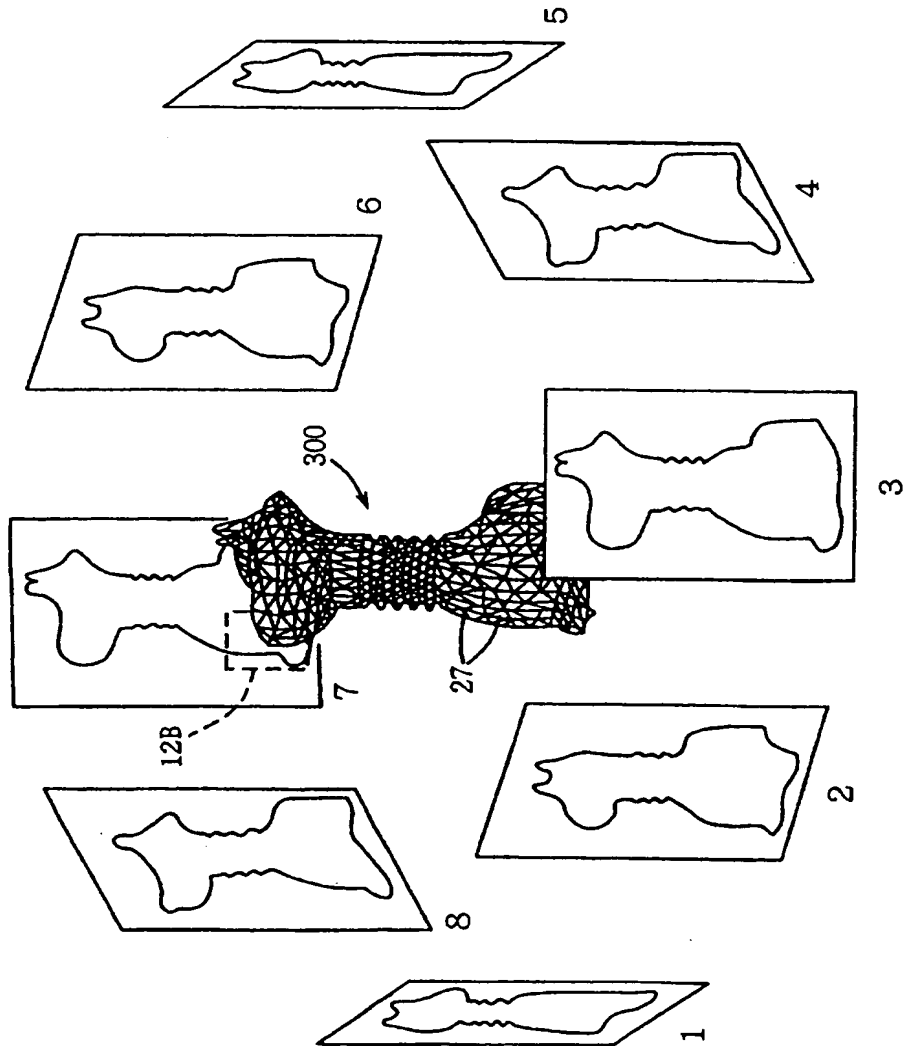




FIG. 13

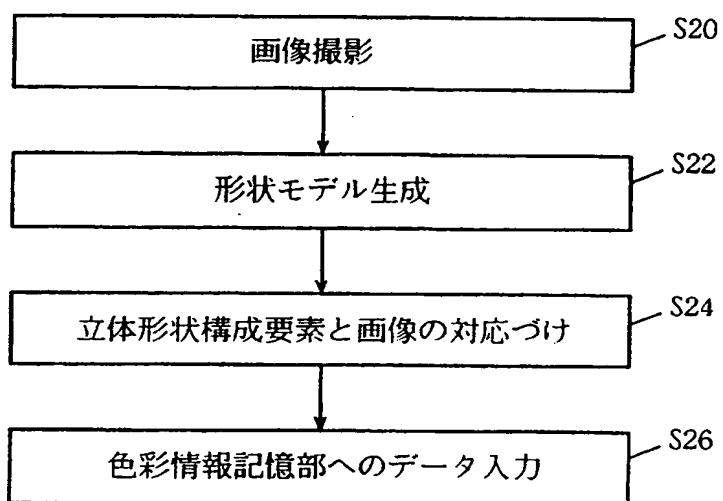


FIG. 14

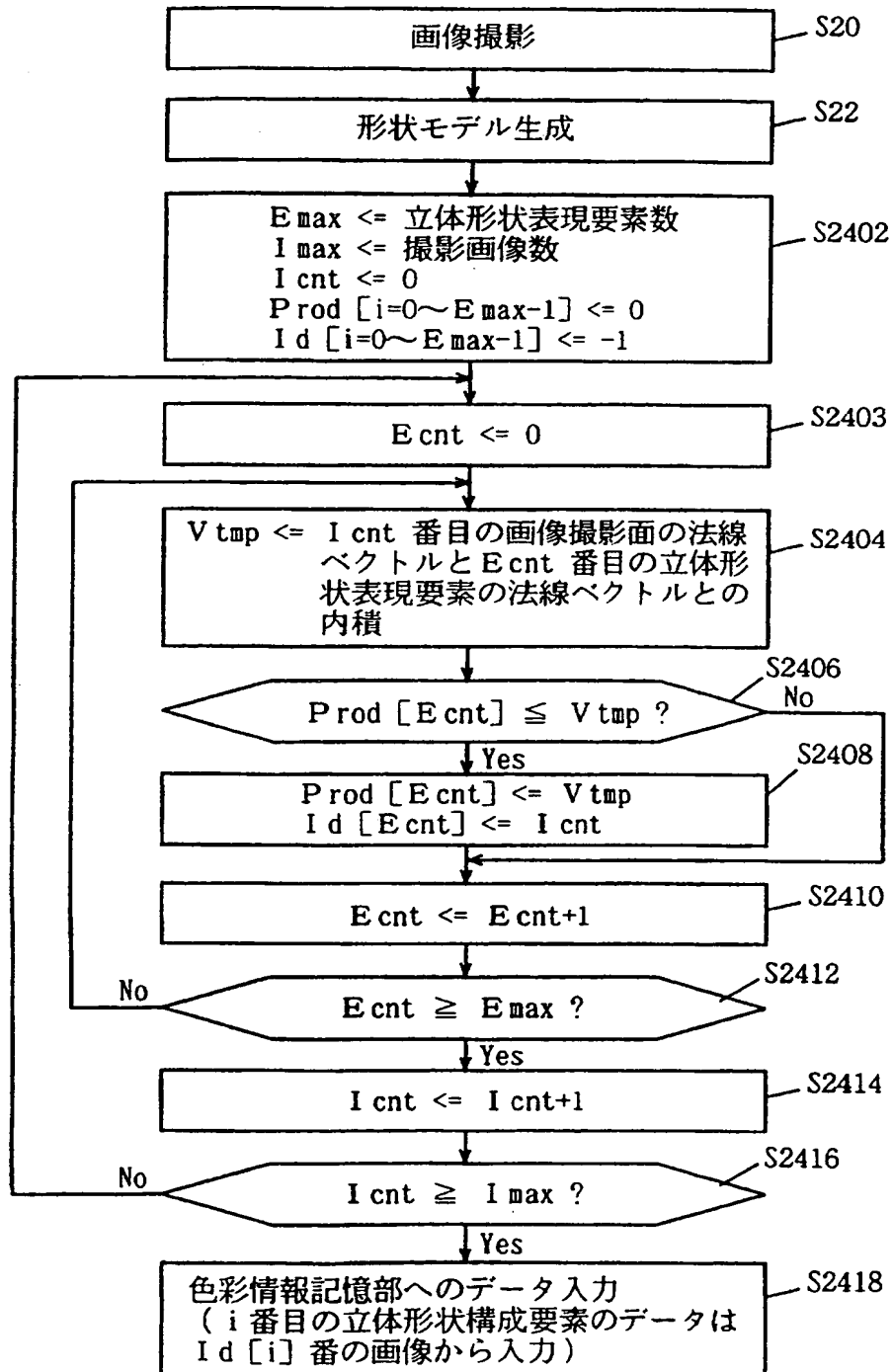


FIG. 15

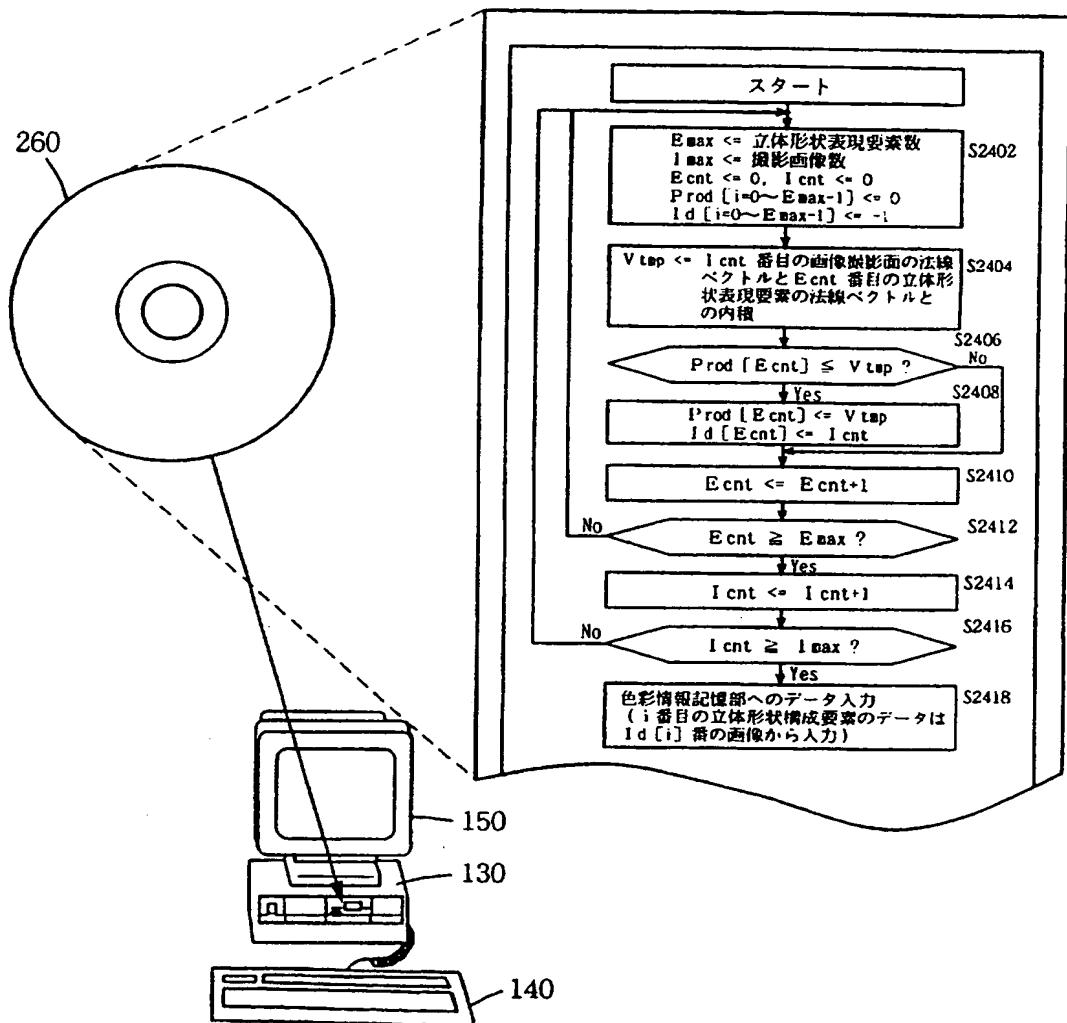


FIG. 16

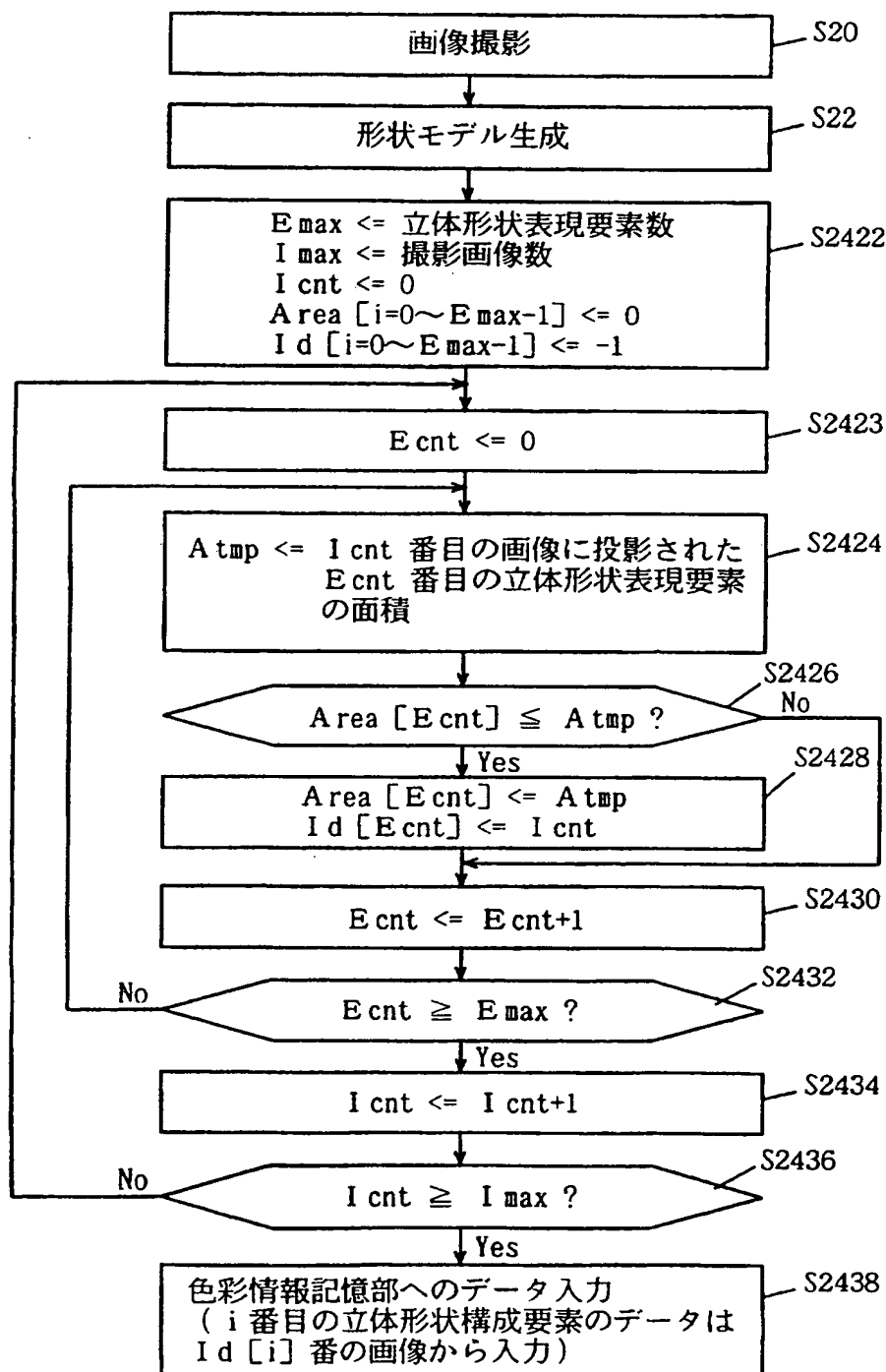


FIG. 17

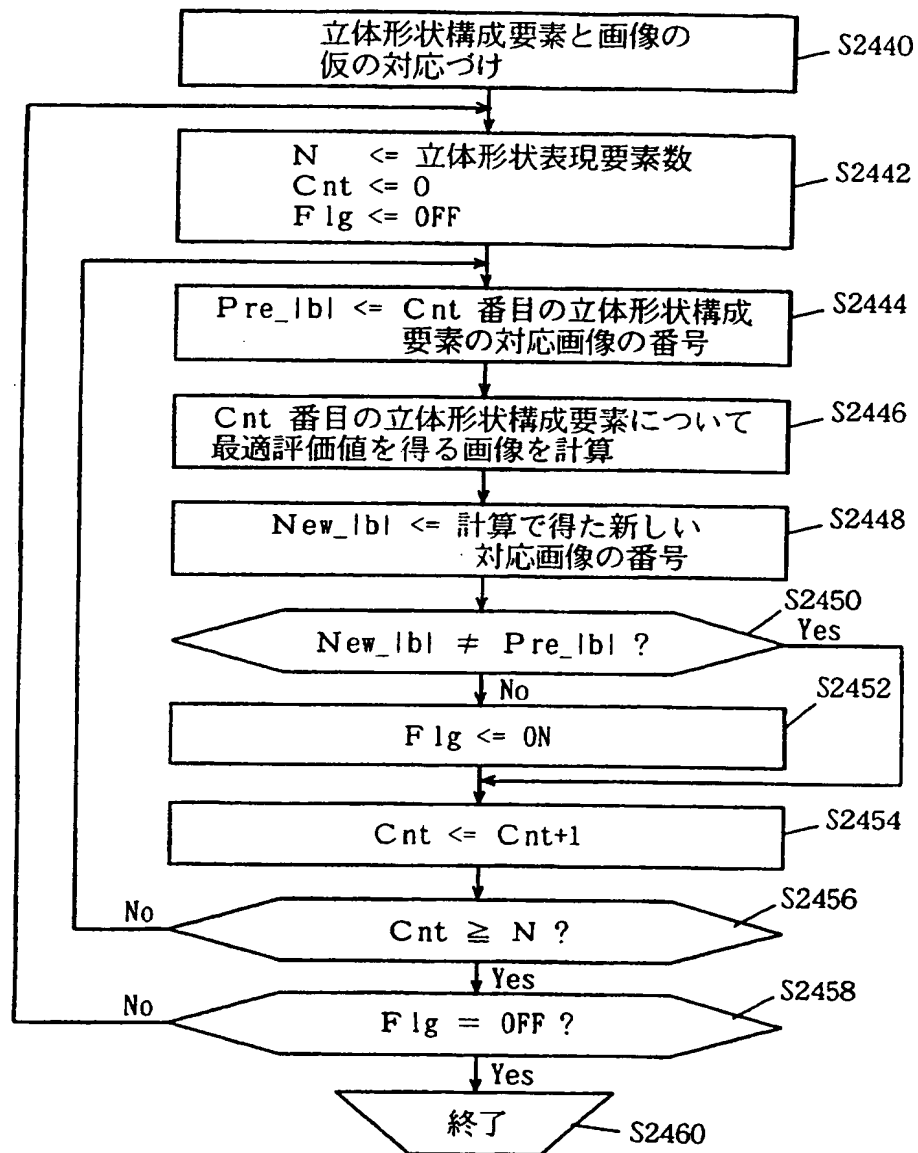


FIG. 18

画像に投影されたポリゴン

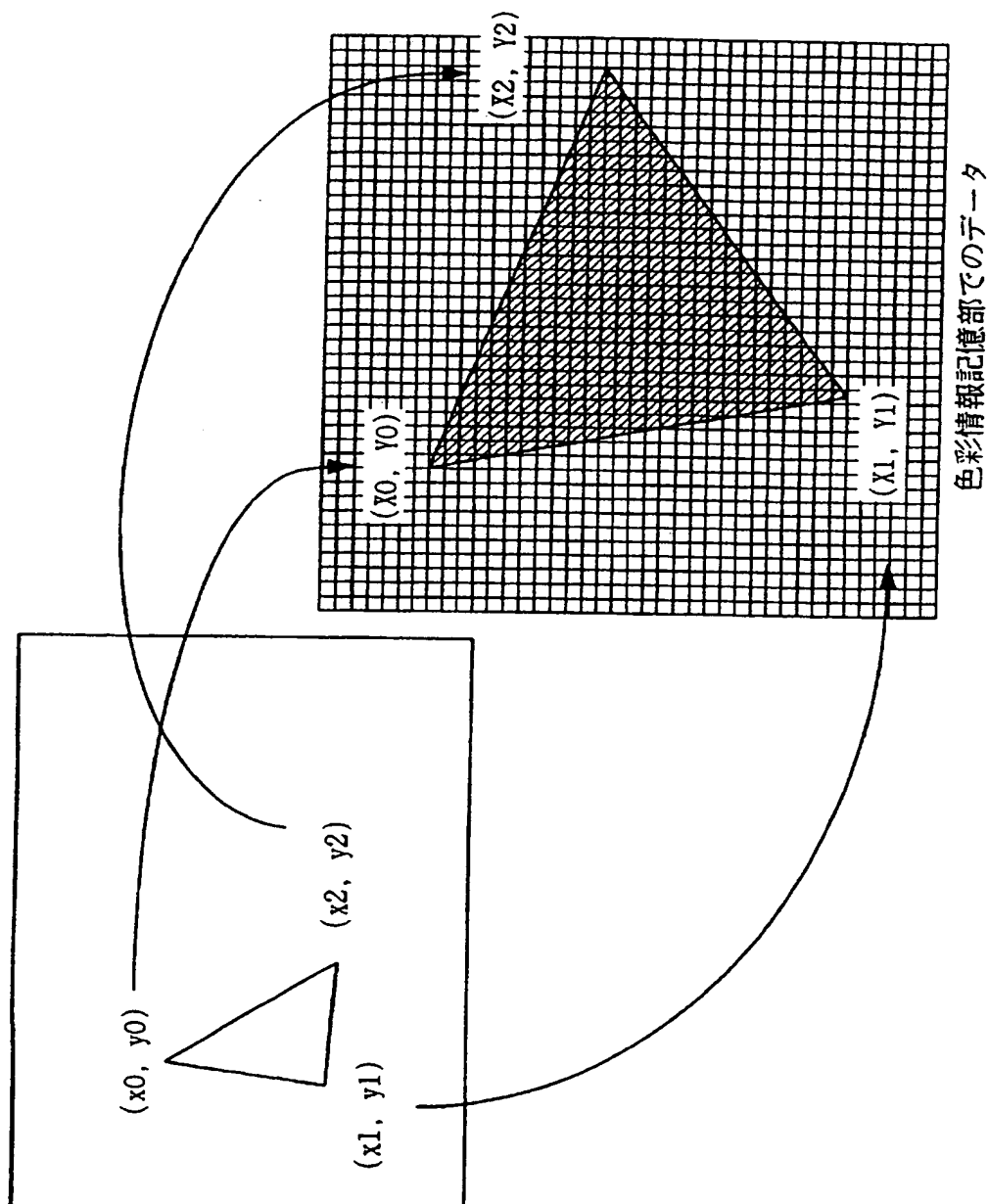


FIG. 19

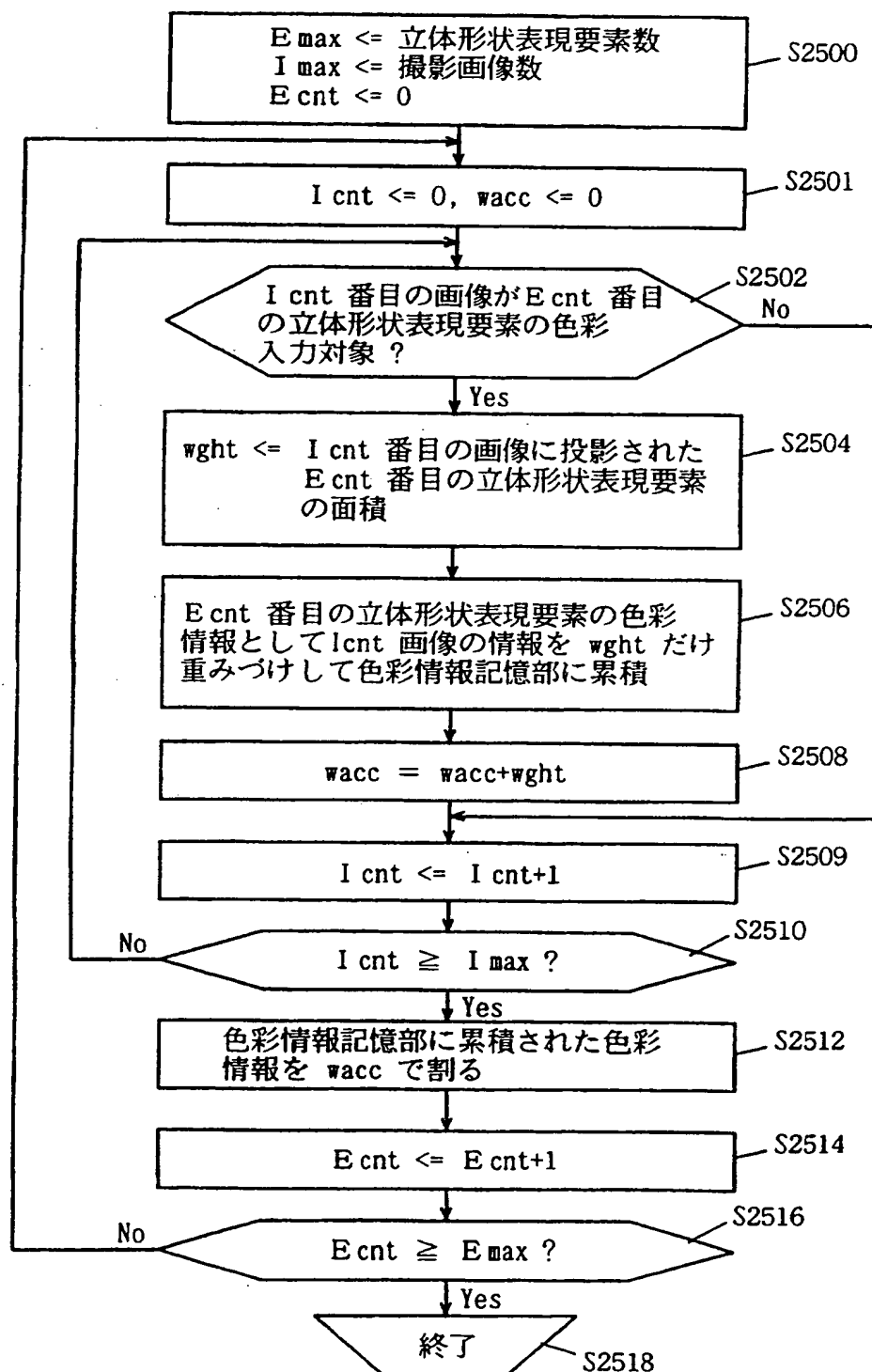


FIG. 20

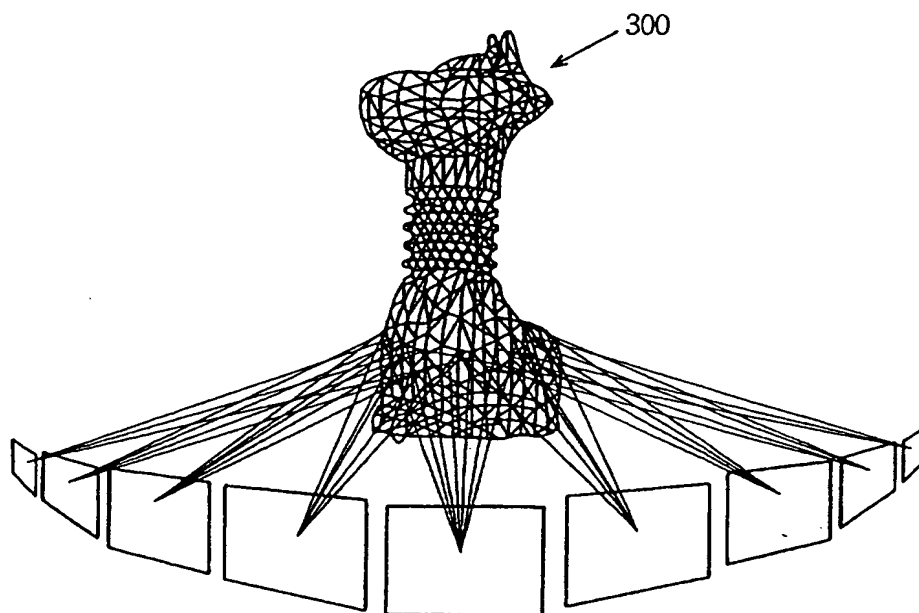


FIG. 21

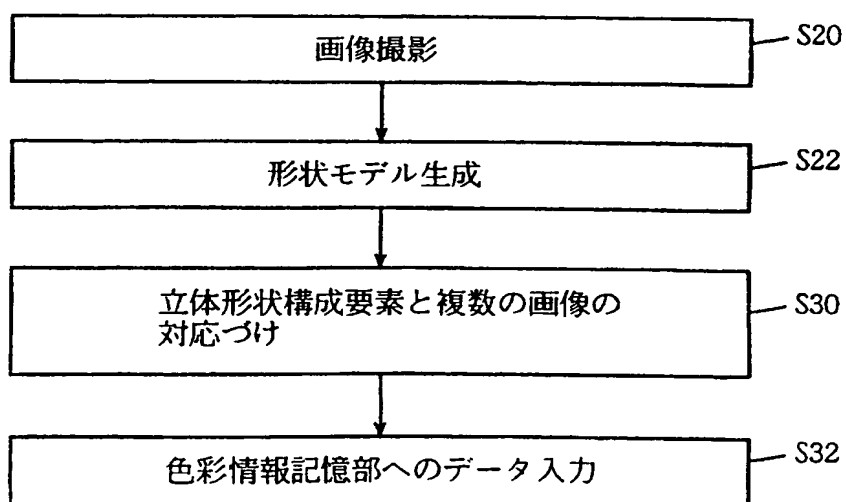




FIG. 22

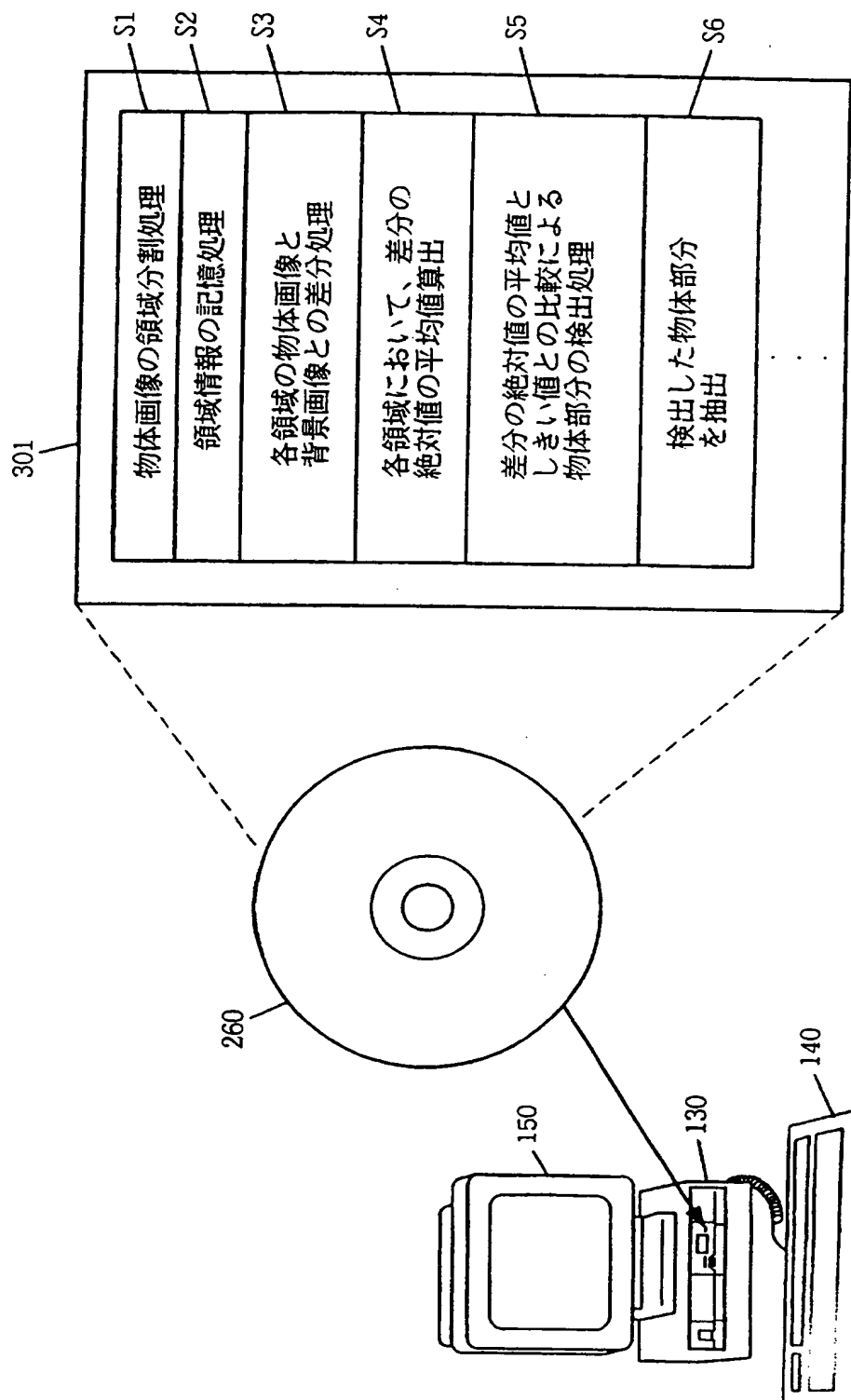


FIG. 23

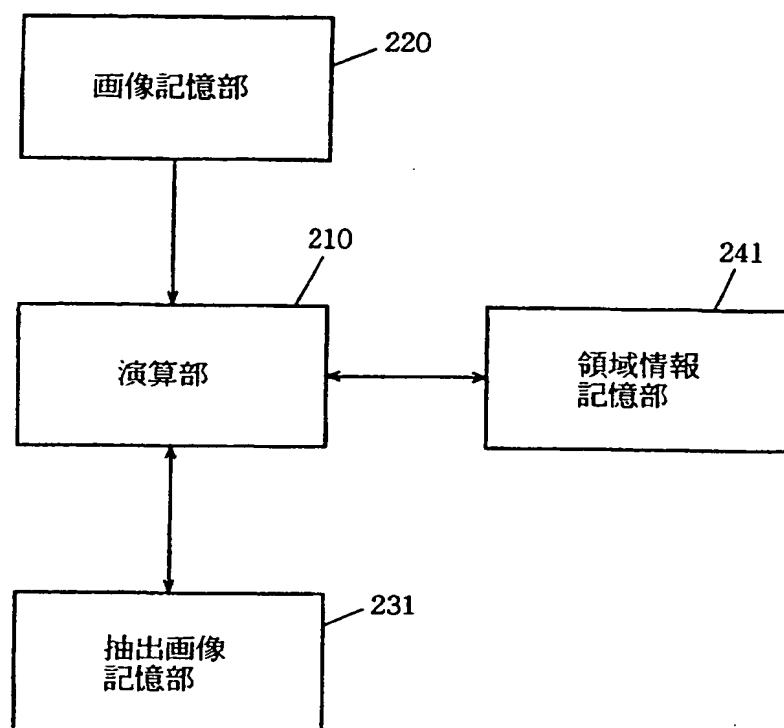


FIG. 24

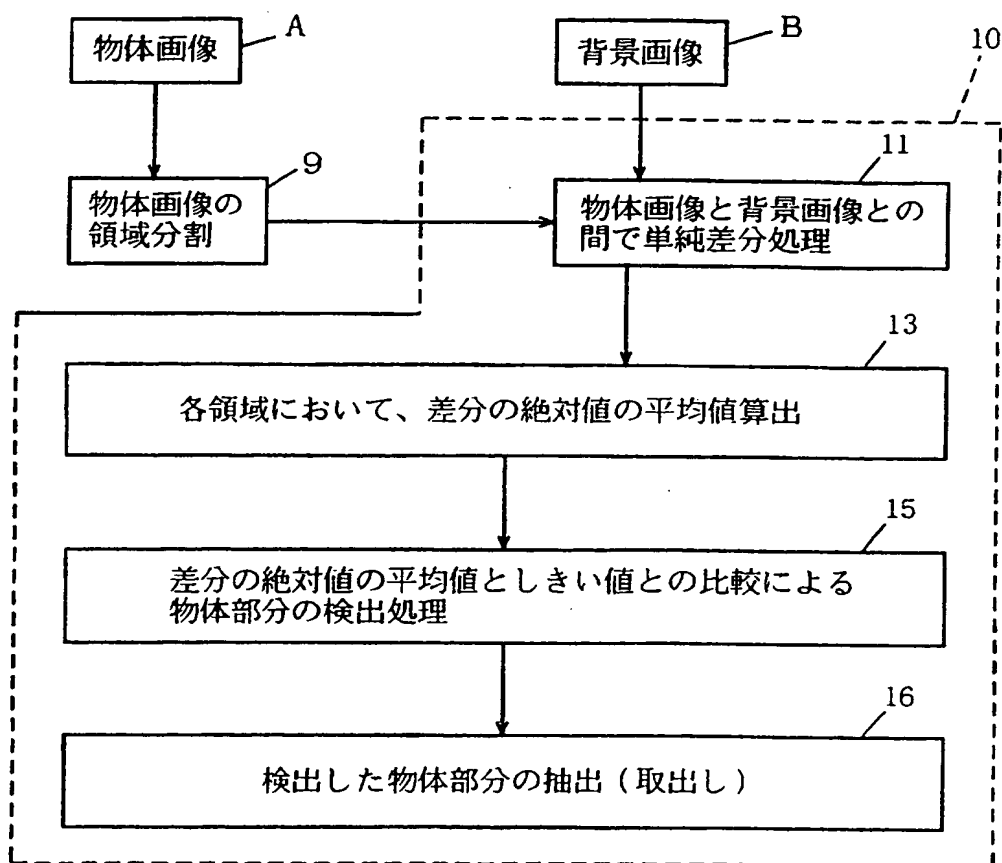


FIG. 25C

FIG. 25B

FIG. 25A

対象物体画像17

背景画像23

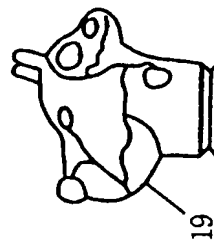
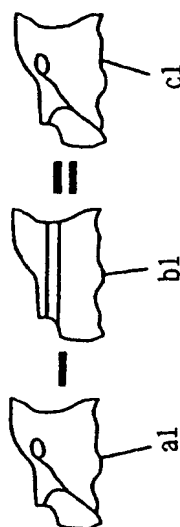
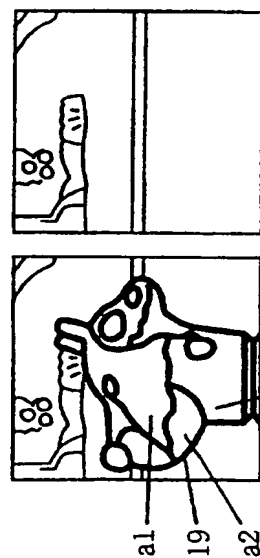


FIG. 26

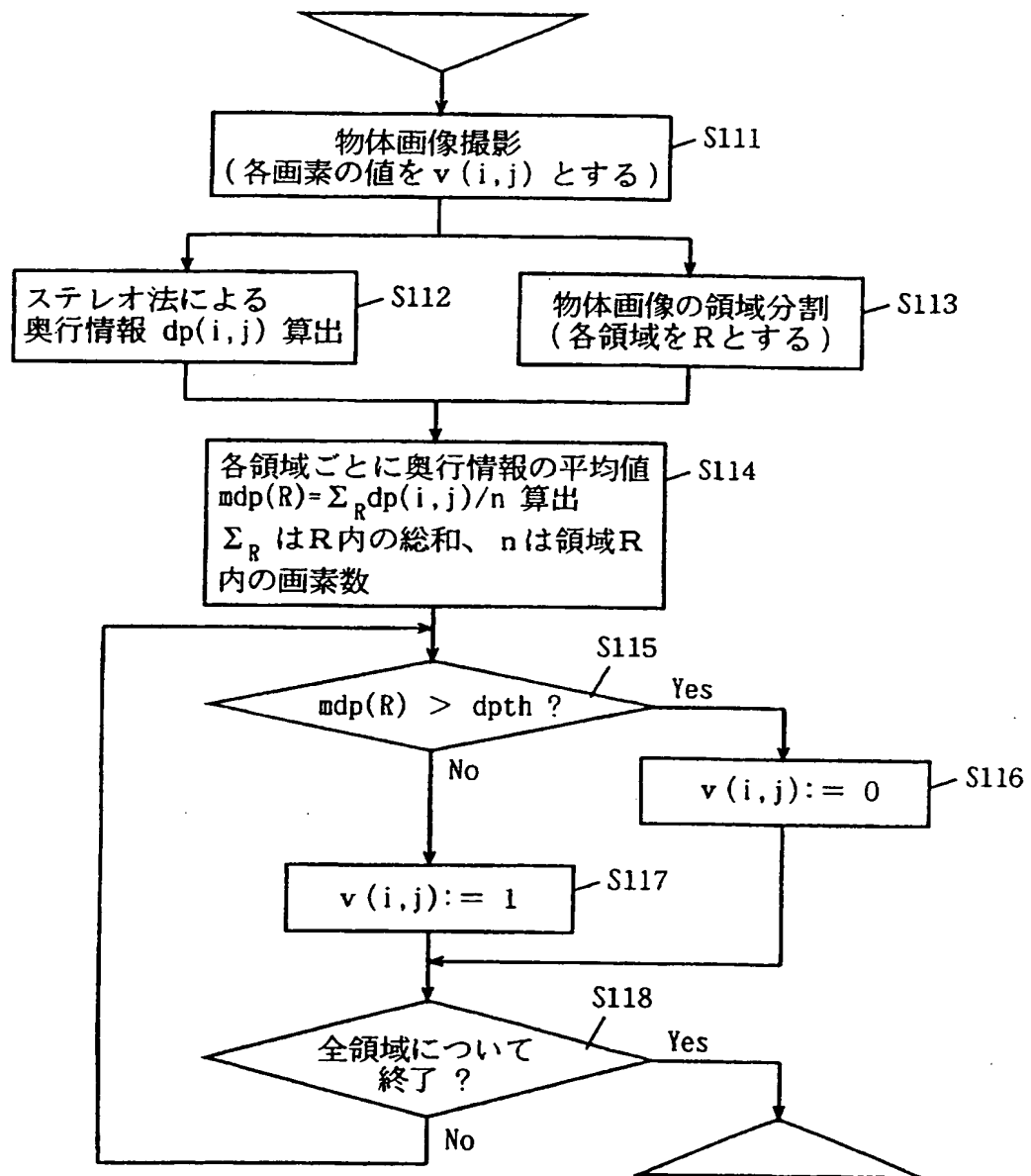
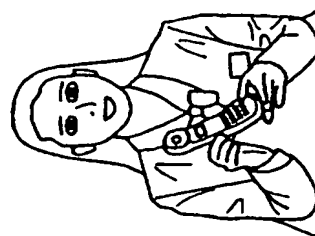
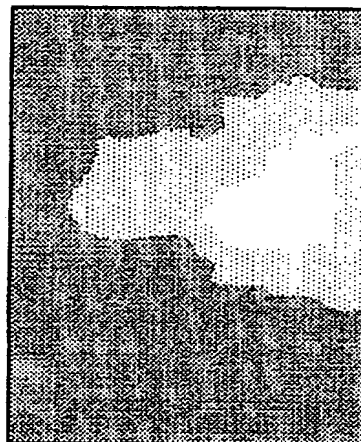


FIG. 27C



切り出し画像

FIG. 27B



奥行情報画像

FIG. 27A



領域分割画像

FIG. 28

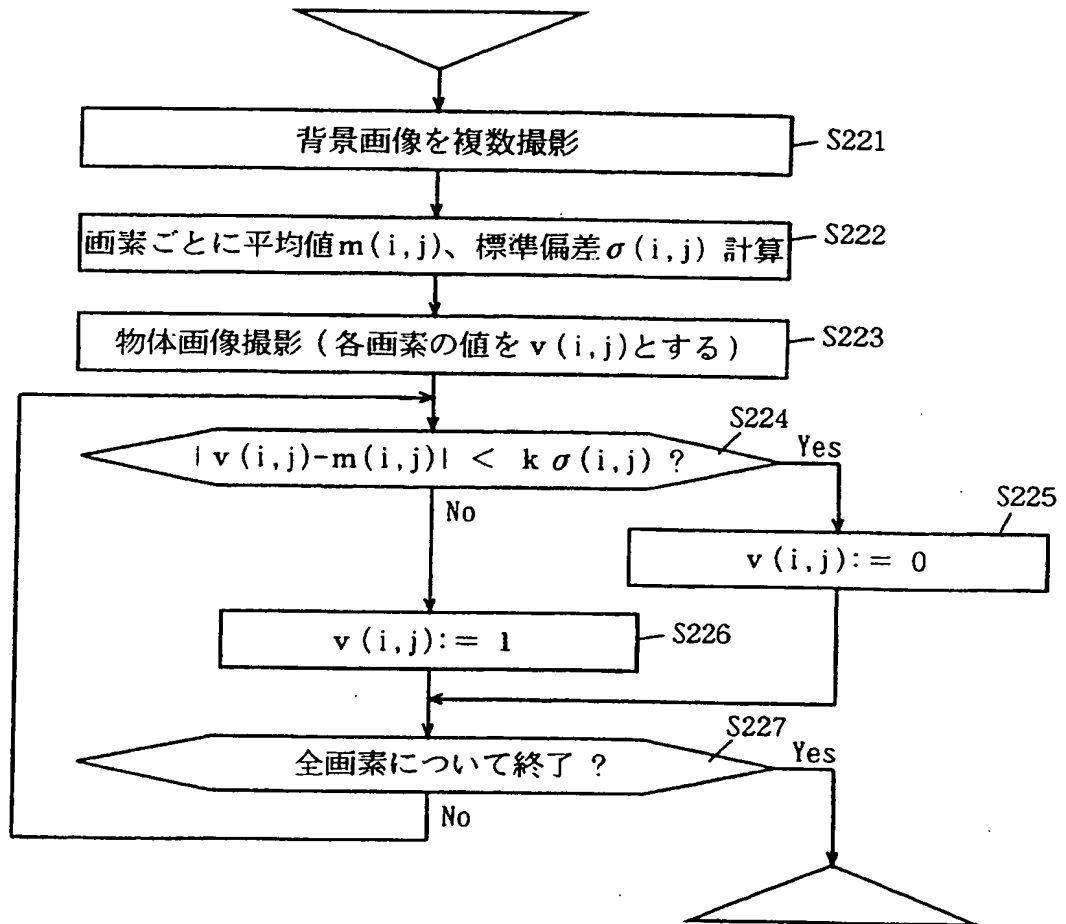


FIG. 29

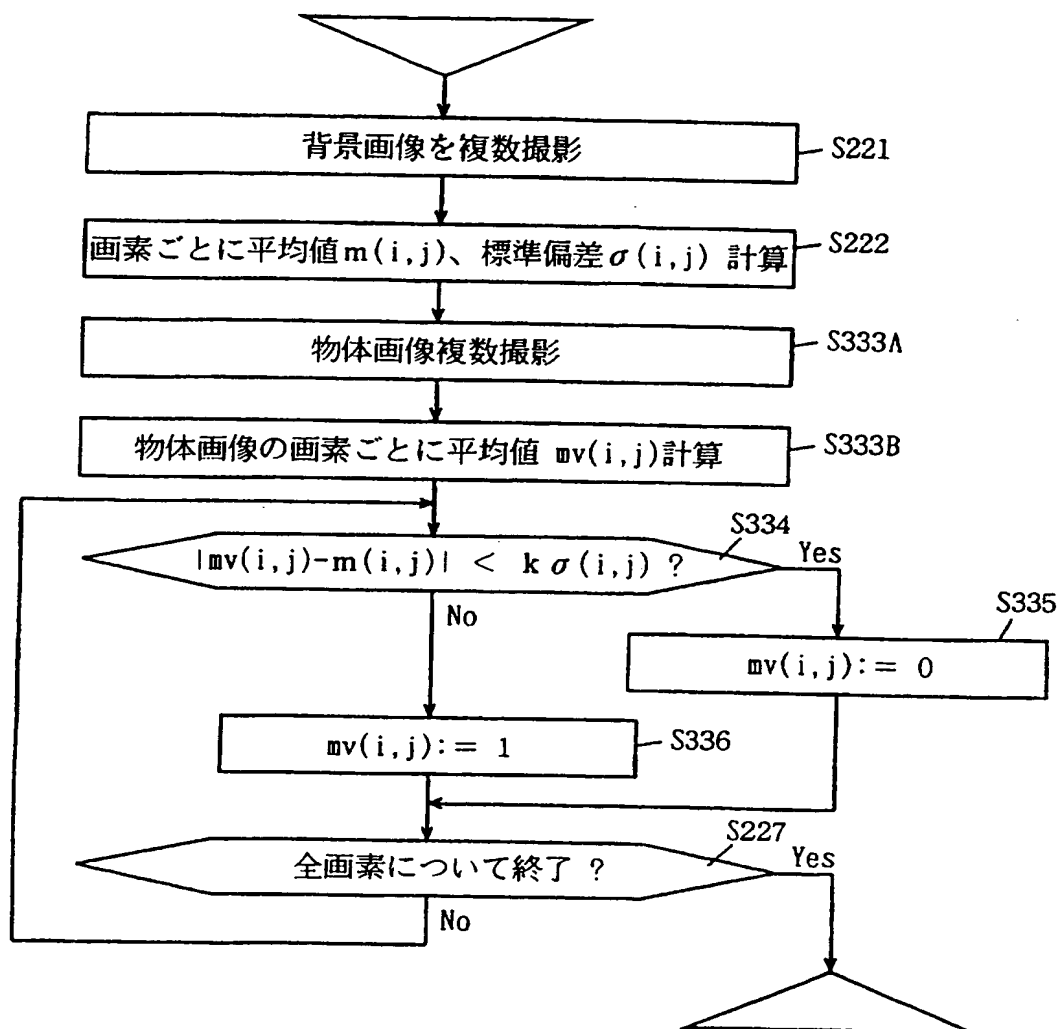




FIG. 30

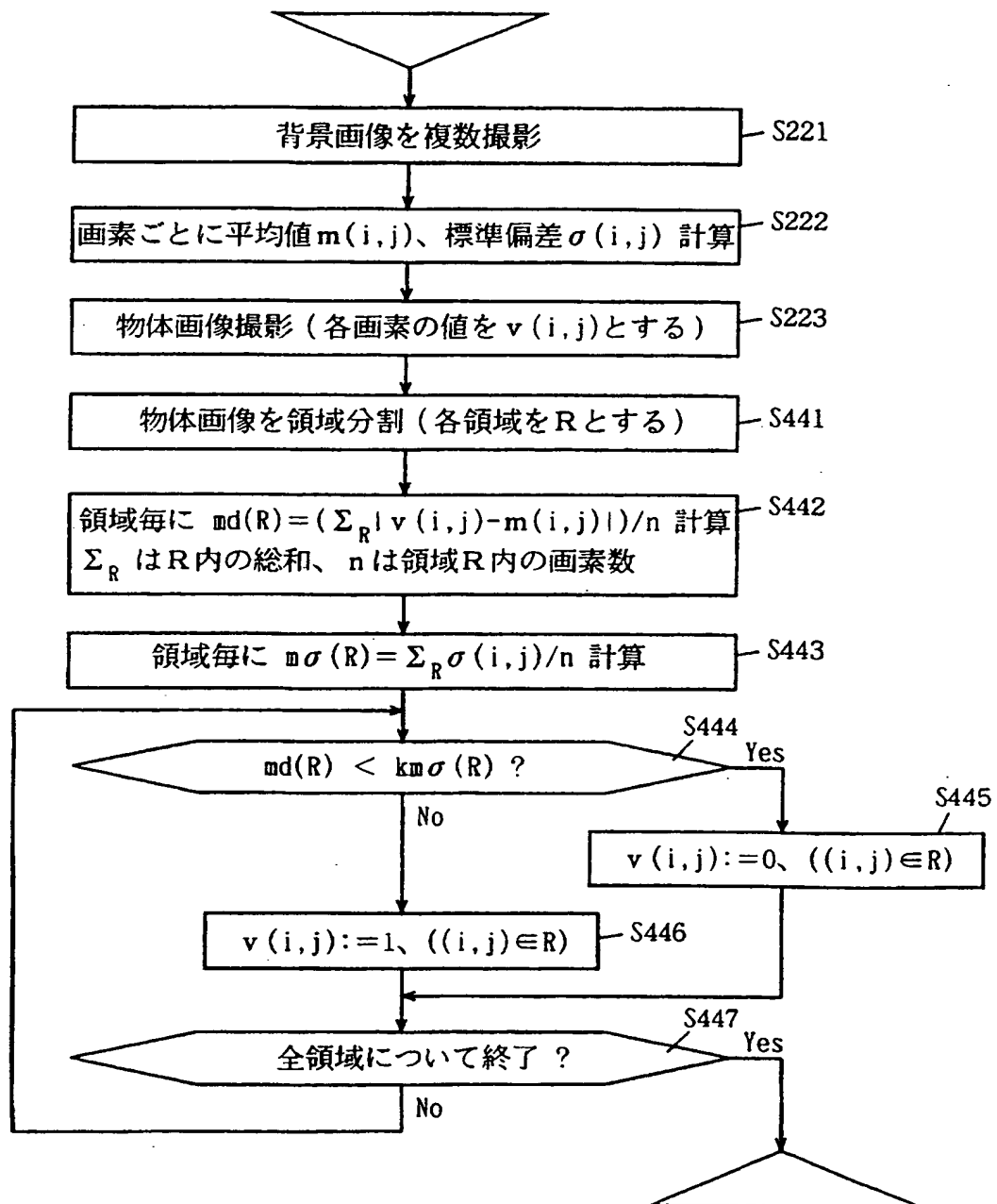


FIG. 31

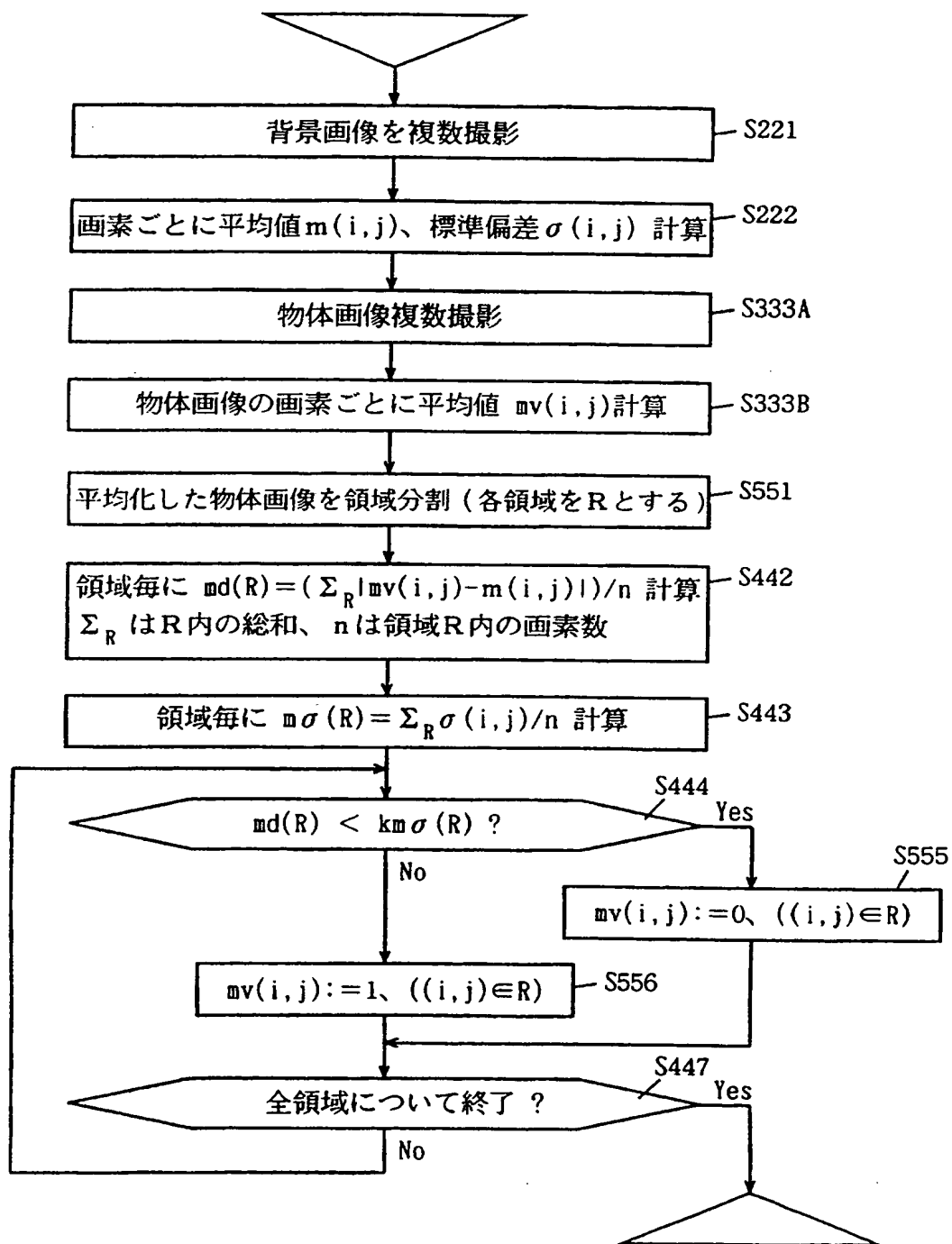


FIG. 32

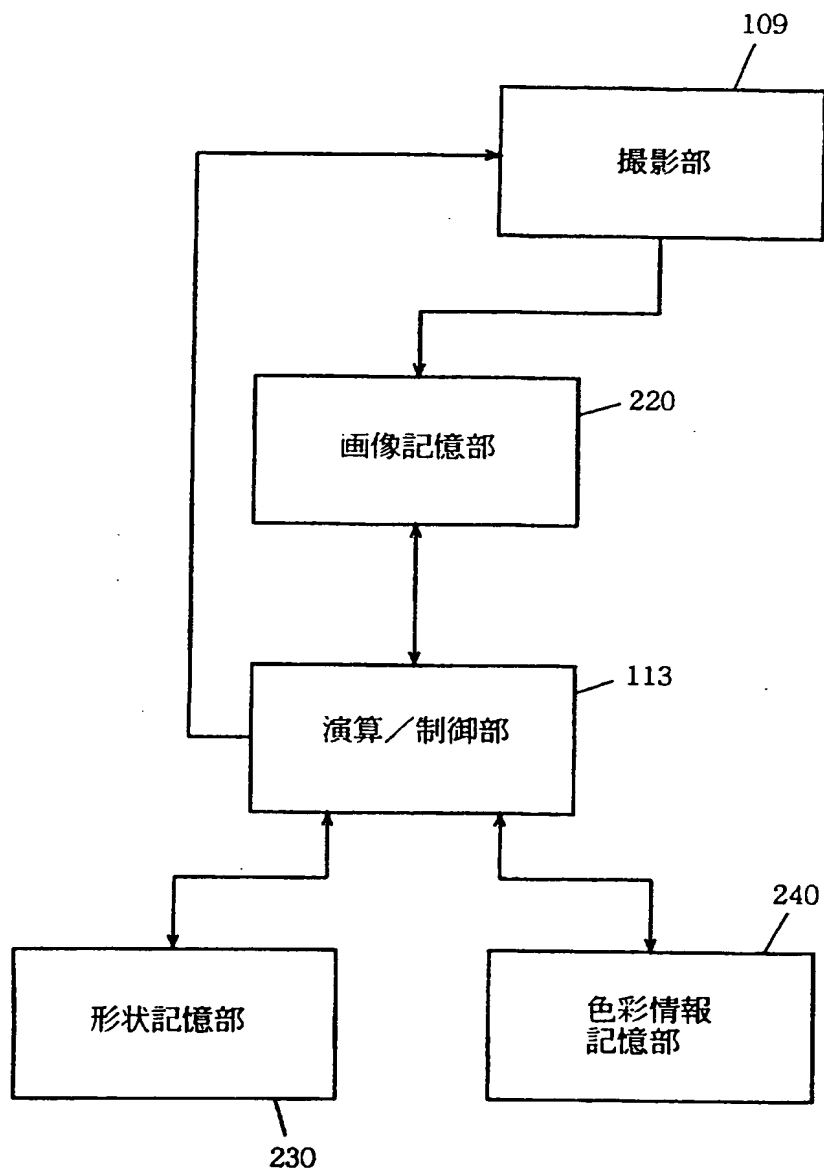


FIG. 33

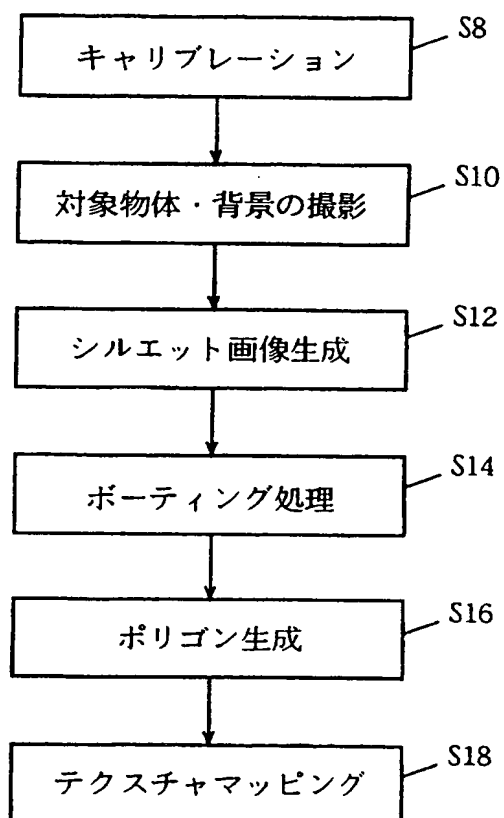


FIG. 34

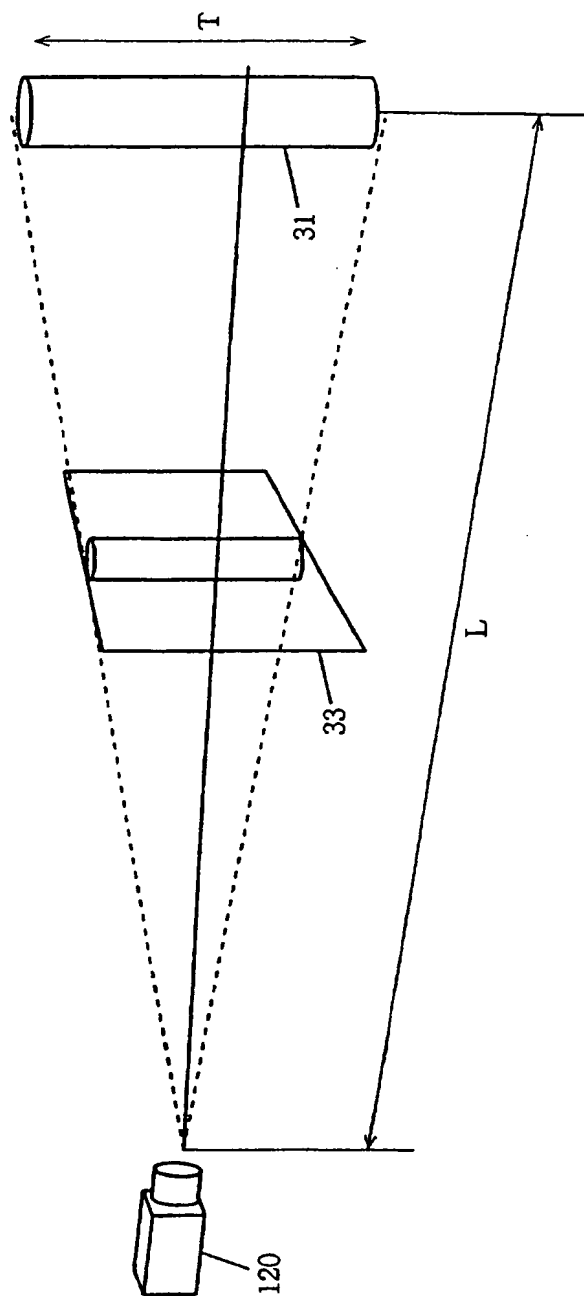


FIG. 35A

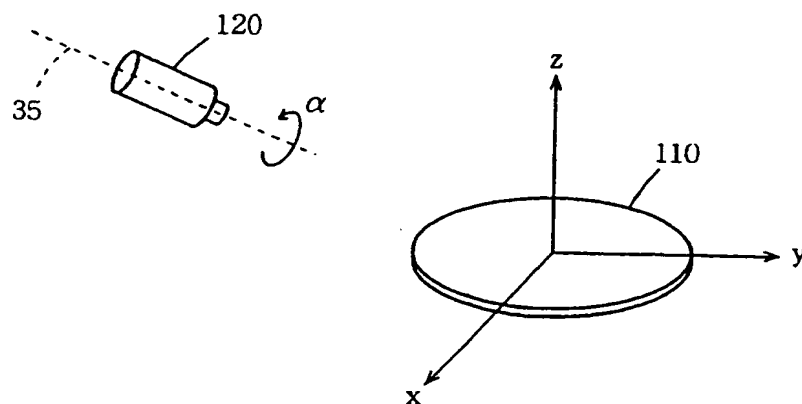


FIG. 35B

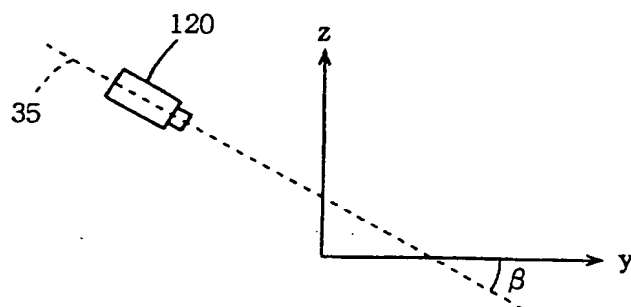


FIG. 35C

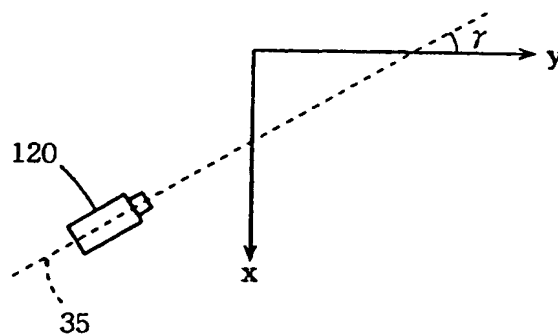


FIG. 36

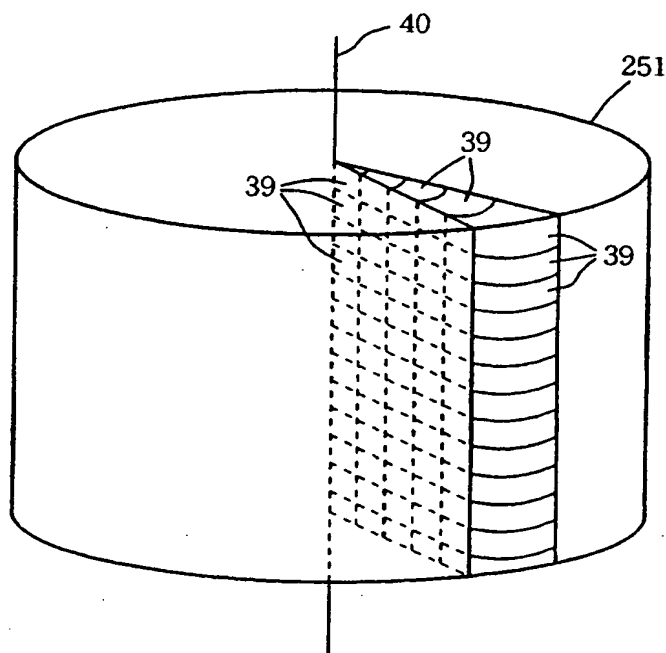


FIG. 37

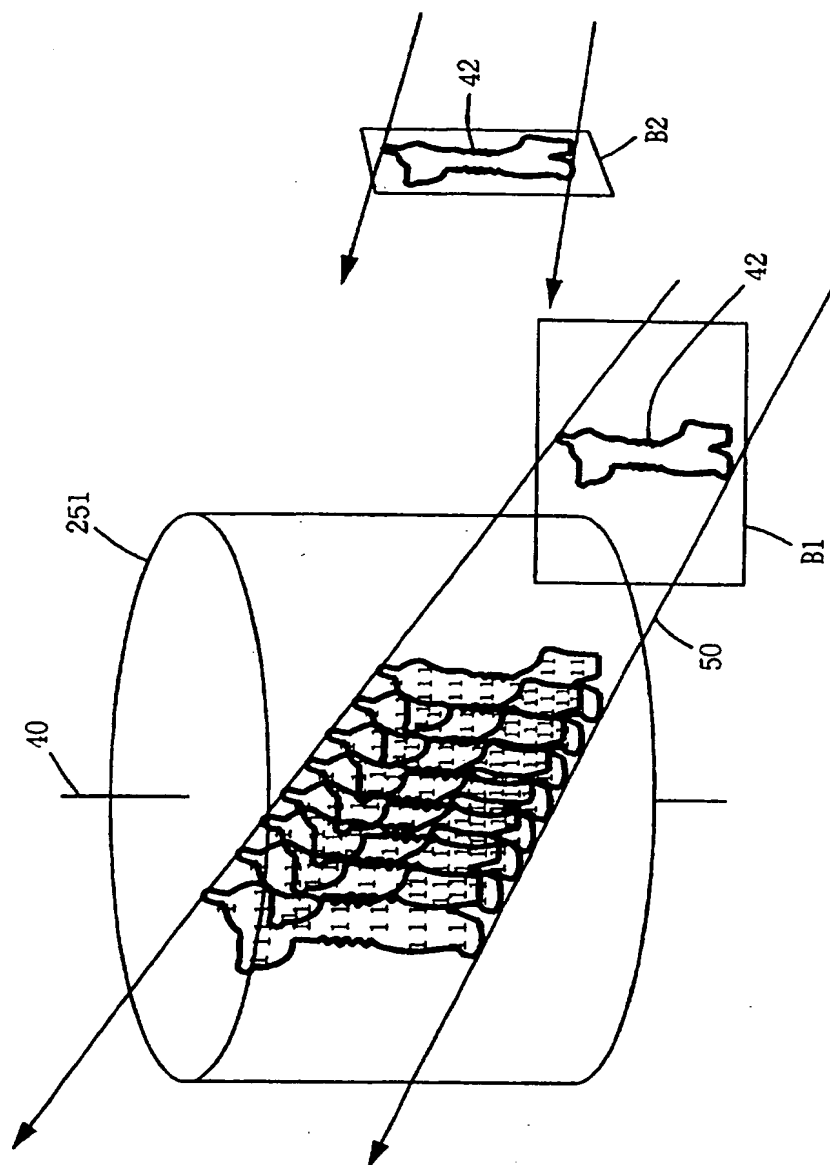




FIG. 38

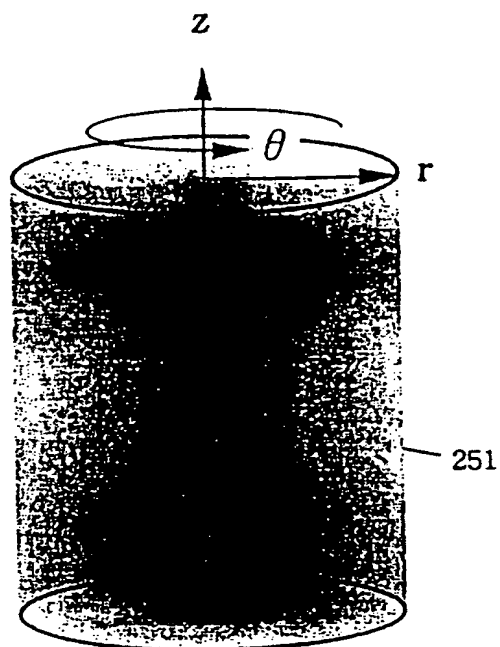


FIG. 39A

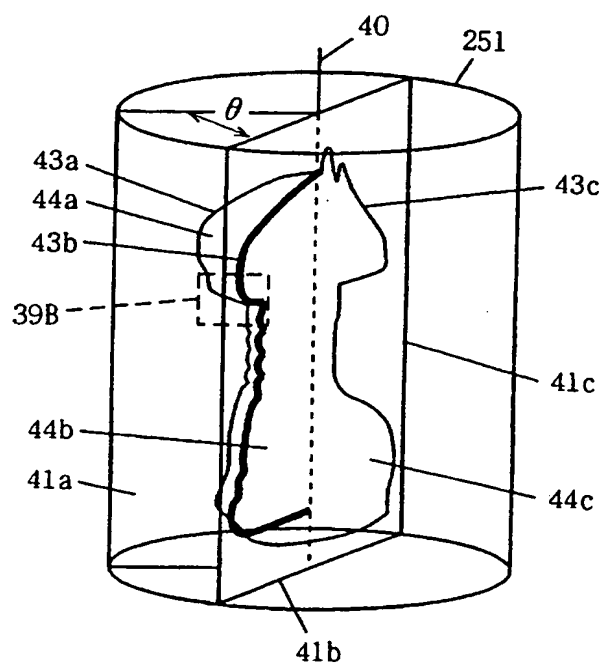


FIG. 39B

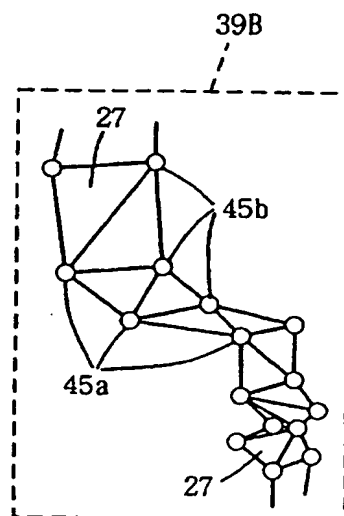


FIG. 40

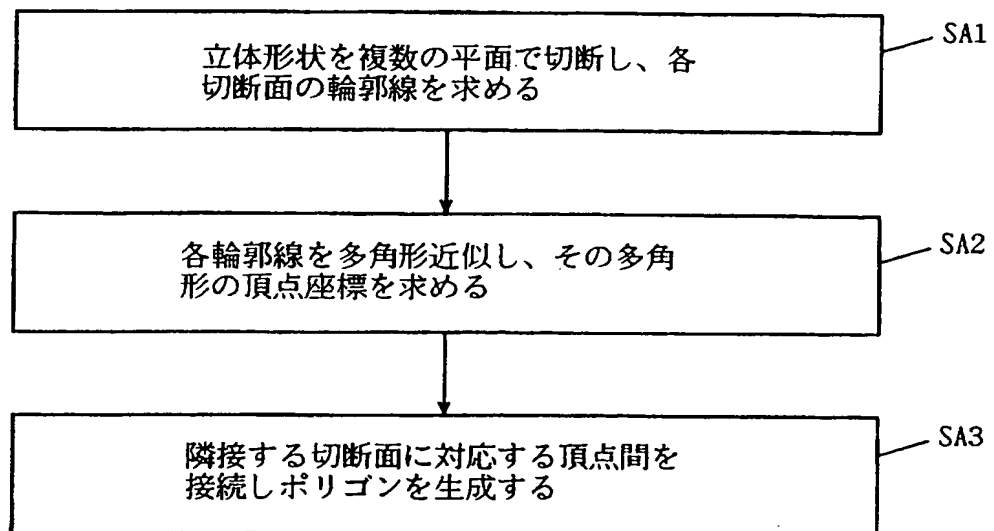


FIG. 41

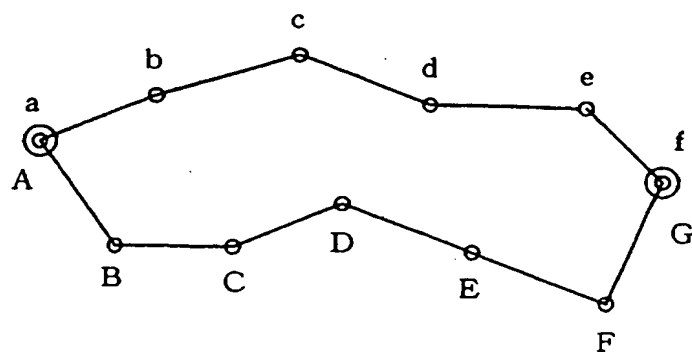


FIG. 42

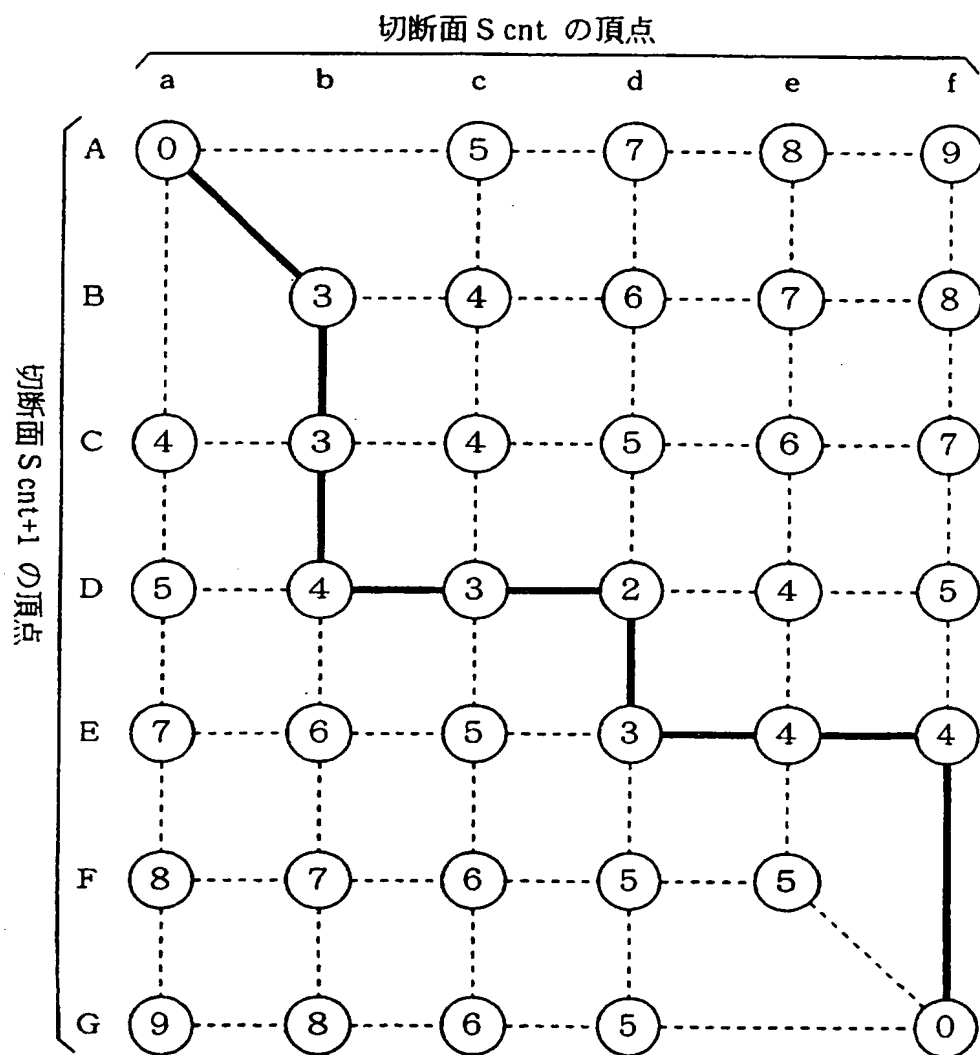


FIG. 43

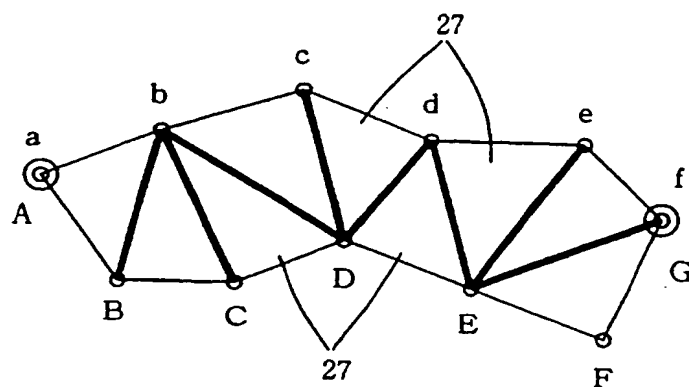


FIG. 44

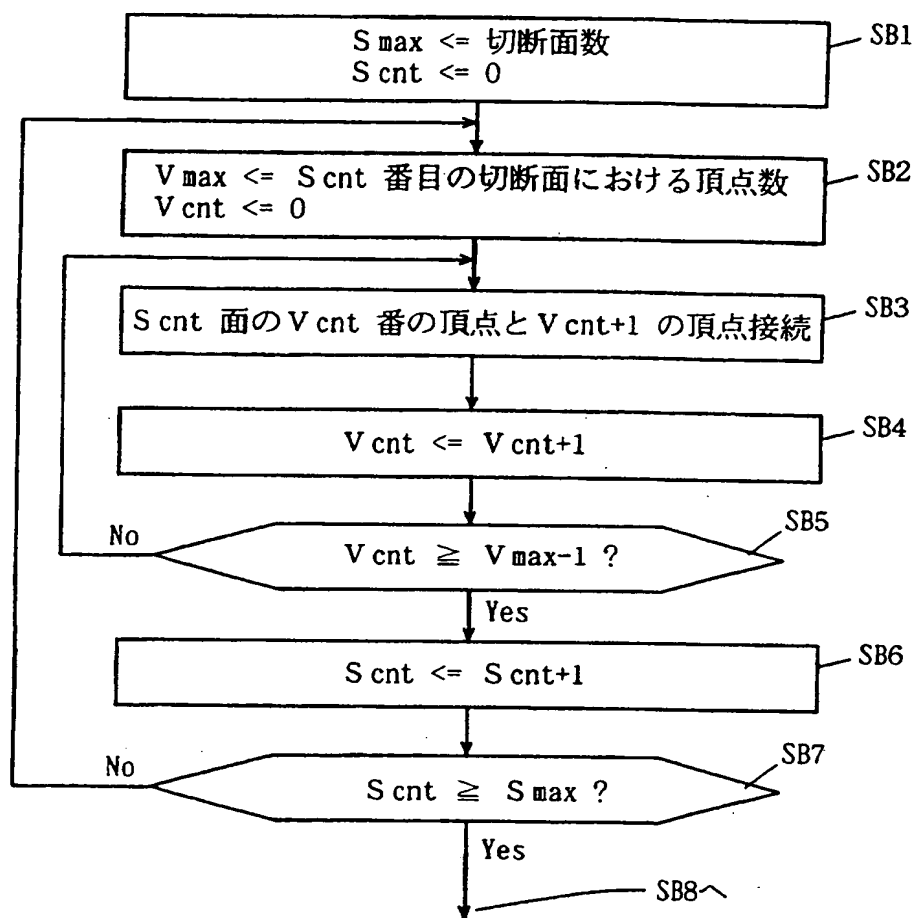


FIG. 45

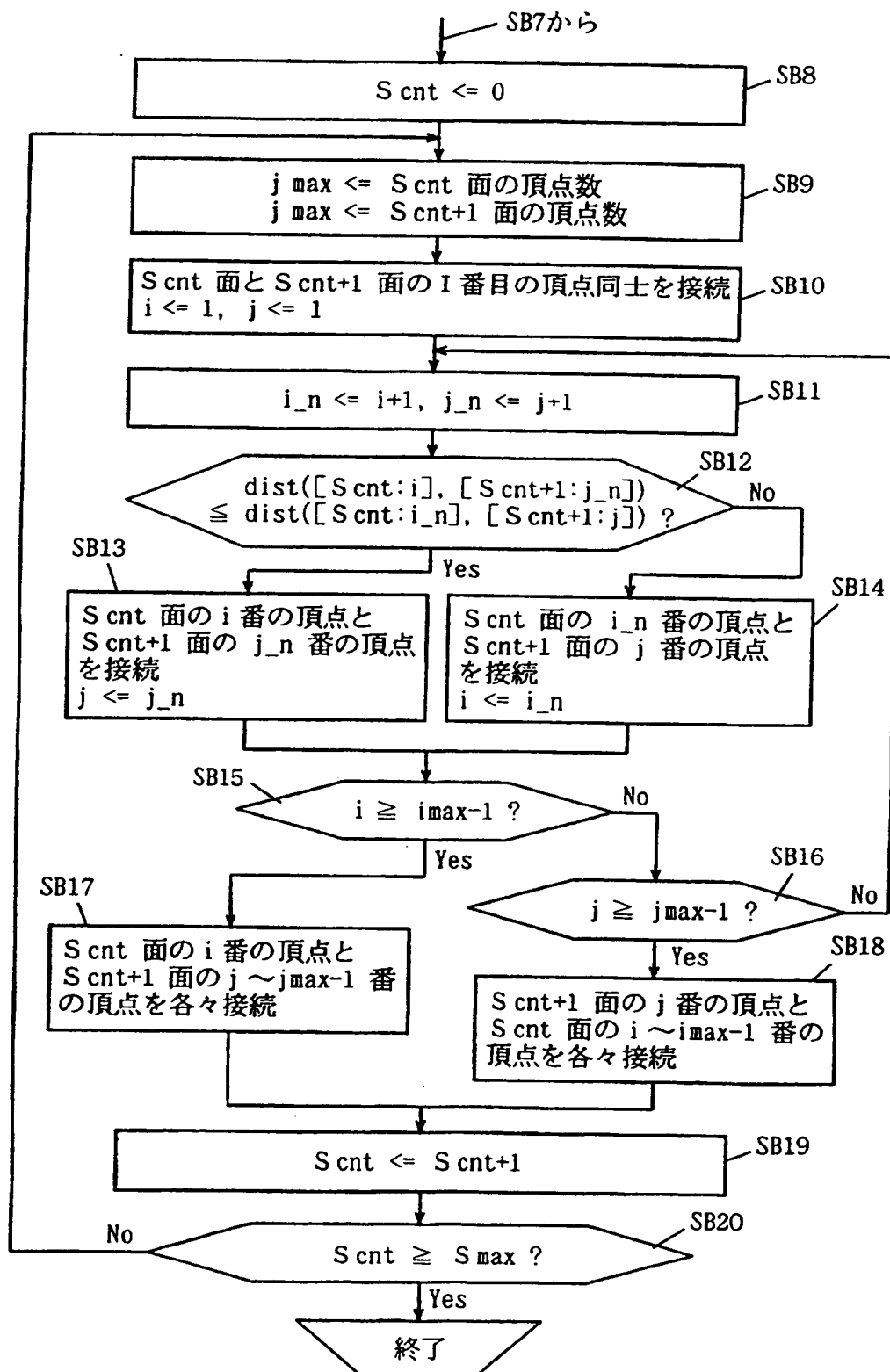




FIG. 46

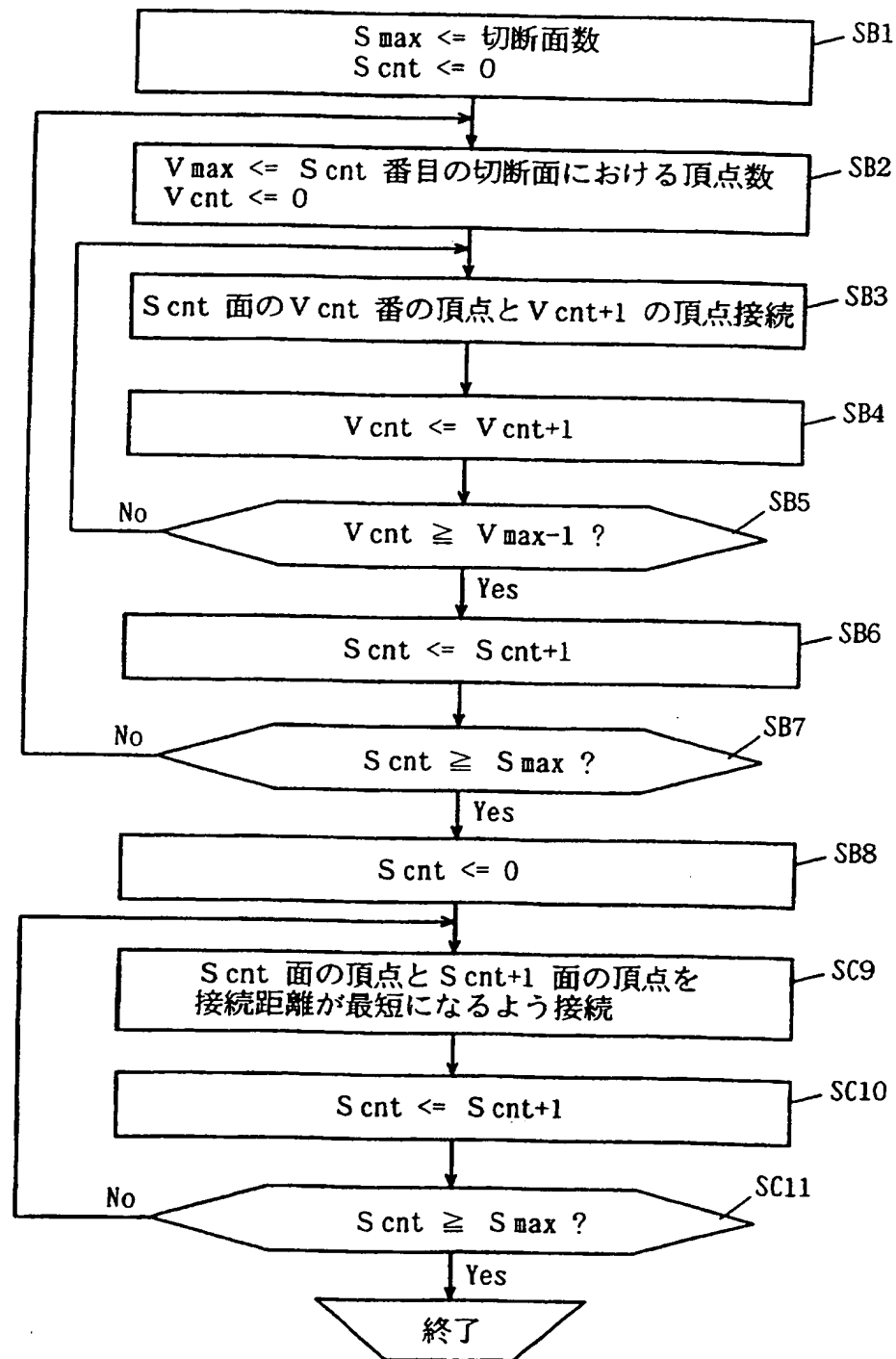
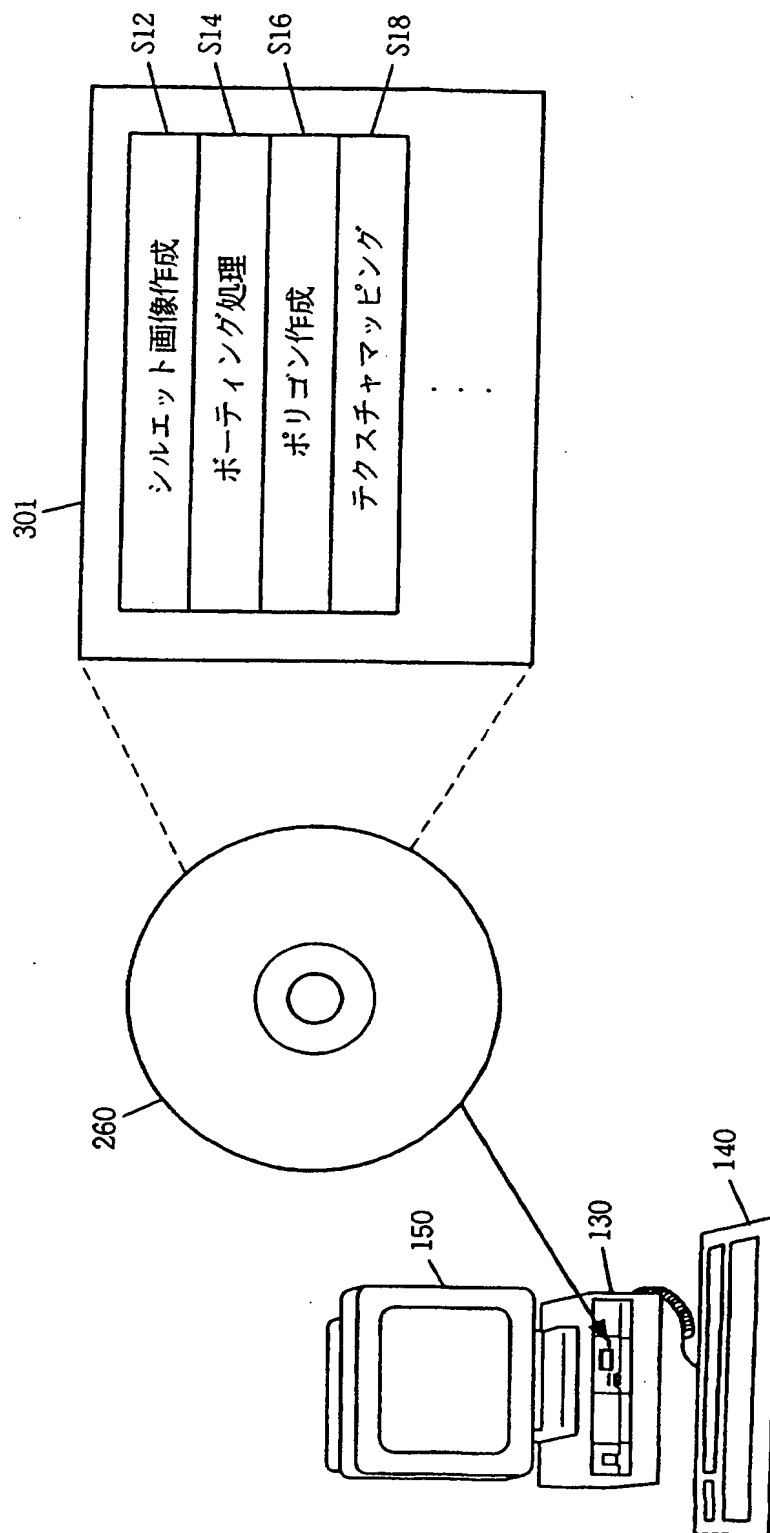


FIG. 47



## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP97/02997

## A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl<sup>6</sup> G06T15/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int. Cl<sup>6</sup> G06T1/00-17/50

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

|                            |             |
|----------------------------|-------------|
| Jitsuyo Shinan Koho        | 1926 - 1996 |
| Kokai Jitsuyo Shinan Koho  | 1971 - 1996 |
| Toroku Jitsuyo Shinan Koho | 1994 - 1997 |

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages                         | Relevant to claim No. |
|-----------|--|-----------------------|
| A         | JP, 6-258048, A (Toshiba Corp.),<br>September 16, 1994 (16. 09. 94) (Family: none)                         | 1 - 97                |
| A         | JP, 8-14858, A (Nippon Hoso Kyokai),<br>January 19, 1996 (19. 01. 96) (Family: none)                       | 1 - 66                |
| A         | JP, 8-147494, A (Matsushita Electric Industrial<br>Co., Ltd.),<br>June 7, 1996 (07. 06. 96) (Family: none) | 1 - 66                |

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&amp;" document member of the same patent family

Date of the actual completion of the international search  
November 17, 1997 (17. 11. 97)Date of mailing of the international search report  
December 2, 1997 (02. 12. 97)Name and mailing address of the ISA/  
Japanese Patent Office  
Facsimile No.

Authorized officer

Telephone No.

## A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. Cl<sup>8</sup> G06T 15/00

## B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int. Cl<sup>8</sup> G06T 1/00-17/50

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1926-1996年

日本国公開実用新案公報 1971-1996年

日本国登録実用新案公報 1994-1997年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

## C. 関連すると認められる文献

| 引用文献の<br>カテゴリー* | 引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示                              | 関連する<br>請求の範囲の番号 |
|-----------------|--|------------------|
| A               | J P, 6-258048, A (株式会社東芝) 16. 9月. 1994 (16. 09. 94) &ファミリーなし   | 1-97             |
| A               | J P, 8-14858, A (日本放送協会) 19. 1月. 1996 (19. 01. 96) &ファミリーなし    | 1-66             |
| A               | J P, 8-147494, A (松下電器産業株式会社) 7. 6月. 1996 (7. 06. 96) &ファミリーなし | 1-66             |

☐ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

## \* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの

「E」先行文献ではあるが、国際出願日以後に公表されたもの

「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)

「O」口頭による開示、使用、展示等に言及する文献

「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&amp;」同一パテントファミリー文献

国際調査を完了した日

17. 11. 97

国際調査報告の発送日

02.12.97

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)

郵便番号100

東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

岡本 俊成

印

5H

9178

電話番号 03-3581-1101 内線 3531



|   |  |  |                                   |
|---|--|--|-----------------------------------|
| (51) 国際特許分類6<br>G06T 15/00  |  | A1   | (11) 国際公開番号<br>WO98/09253         |
|   |  |  | (43) 国際公開日<br>1998年3月5日(05.03.98) |
| (21) 国際出願番号<br>PCT/JP97/02997   |  | (74) 代理人<br>荒川 勉(ARAKAWA, Tsutomu)[JP/JP]<br>〒301 茨城県竜ヶ崎市長山6-15-19 Ibaraki, (JP)                   |                                   |
| (22) 国際出願日<br>1997年8月28日(28.08.97)  |  | (74) 代理人<br>弁理士 深見久郎, 外(FUKAMI, Hisao et al.)<br>〒530 大阪府大阪市北区南森町2丁目1番29号<br>住友銀行南森町ビル Osaka, (JP) |                                   |
| (30) 優先権データ<br>特願平8/248739 1996年8月29日(29.08.96) JP<br>特願平8/248946 1996年8月30日(30.08.96) JP<br>特願平8/248958 1996年8月30日(30.08.96) JP<br>特願平9/214104 1997年7月23日(23.07.97) JP<br>特願平9/217486 1997年8月12日(12.08.97) JP<br>特願平9/217550 1997年8月12日(12.08.97) JP   |  | (81) 指定国 CN, US.<br>添付公開書類<br>国際調査報告書  |                                   |
| (71) 出願人 (米国を除くすべての指定国について)<br>三洋電機株式会社(SANYO ELECTRIC CO., LTD.)(JP/JP)<br>〒570 大阪府守口市京阪本通2丁目5番5号 Osaka, (JP)  |  |  |                                   |
| (72) 発明者; および<br>(75) 発明者/出願人 (米国についてののみ)<br>松本幸則(MATSUMOTO, Yukinori)[JP/JP]<br>〒305 茨城県つくば市竹園2-16-12-25 Ibaraki, (JP)<br>寺崎 肇(TERASAKI, Hajime)[JP/JP]<br>〒305 茨城県つくば市二の宮4-13-1-2-108 Ibaraki, (JP)<br>杉本和英(SUGIMOTO, Kazuhide)[JP/JP]<br>〒305 茨城県つくば市竹園2-16-12-45 Ibaraki, (JP)  |  |  |                                   |
| (54)Title: TEXTURE INFORMATION GIVING METHOD, OBJECT EXTRACTING METHOD, THREE-DIMENSIONAL MODEL GENERATING METHOD AND APPARATUS FOR THE SAME  |  |  |                                   |
| (54)発明の名称 テクスチャ情報付与方法、物体抽出方法、立体モデル作成方法、およびその装置  |  |  |                                   |
| (57) Abstract<br>This method is such that a three-dimensional model of an actual object is expressed by polygons in accordance with information about a plurality of images formed by imaging the actual object while the object is turned in steps of an arbitrary angle, and texture information is given to each polygon from information about the object image whose polygon has the largest projection area. In order to improve the color continuity between adjacent polygons, the object image information corresponding to the polygon concerned and to the polygon adjacent to it are so selected as to be the image information about object images which are picked up from directions close to each other in positions close to each other. Another method is such that an object image is divided into a plurality of regions, the differentials between the object image and the background image is obtained for each region, the average of the absolute values of the differentials is obtained for each region and regions whose averages of the absolute values of the differentials are larger than a threshold value are detected as object parts. Further, another method is such that only the background of an object is imaged, the object is imaged while it is turned to obtain a plurality of object images, a silhouette image is generated by a differential processing of the object images and the background image, a voting processing of a voxel space is carried out in accordance with the silhouette image, a polygon is generated in accordance with a three-dimensional form obtained by the voting processing, and a texture obtained from the object images is mapped on the polygon. |  |  |                                   |
|   |  |  |                                   |

この方法は、現実の物体を任意の角度ごとに回転させて撮影された複数の物体画像情報に基づいて、3次元形状モデルをポリゴンにより表現し、各ポリゴンに対して当該ポリゴンが最も大きな投影面積を有する物体画像情報からテクスチャ情報を付与する。隣接するポリゴン間の色彩連続性を向上させるために、注目するポリゴンと隣接するポリゴンとに対応付けられる物体画像情報は、撮影位置および撮影方向が近接した物体画像情報となるように選択される。もう1つの方法は、物体画像を複数の領域に分割し、領域単位で物体画像と背景画像との差分を求め、領域単位で差分の絶対値の平均値を導出し、差分の絶対値の平均値がしきい値以上の領域を物体部分として検出する。さらにもう1つの方法は、対象物体の背景だけを撮影するとともに、対象物体を回転させながら撮影して複数の物体画像を得る。そして、物体画像と背景画像との間で差分処理を行なうことによってシルエット画像を生成し、そのシルエット画像をもとにボクセル空間にボーディング処理を行ない、ボーディング処理によって得られた立体形状をもとにポリゴンを生成し、物体画像から獲得したテクスチャをポリゴンにマップする。

PCTに基づいて公開される国際出願のパンフレット第一頁に記載されたPCT加盟国を同定するために使用されるコード (参考情報)

|    |              |    |             |    |                   |    |            |
|----|--------------|----|-------------|----|-------------------|----|------------|
| AL | アルバニア        | ES | スペイン        | LK | スリランカ             | SE | スウェーデン     |
| AM | アルメニア        | FI | フィンランド      | LR | リベリア              | SG | シンガポール     |
| AT | オーストリア       | FR | フランス        | LS | レソト               | SI | スロベニア共和国   |
| AU | オーストラリア      | GA | ガボン         | LT | リトアニア             | SK | スロヴァキア共和国  |
| AZ | アゼルバイジャン     | GB | 英国          | LU | ルクセンブルグ           | SL | シエラレオネ     |
| BA | ボスニア・ヘルツェゴビナ | GE | グルジア        | LV | ラトヴィア             | SN | セネガル       |
| BB | バルバドス        | GH | ガーナ         | MC | モナコ               | SZ | スワジランド     |
| BE | ベルギー         | GM | ガンビア        | MD | モルドヴァ共和国          | TD | チャド        |
| BF | ブルキナ・ファソ     | GN | ギニア         | MG | マダガスカル            | TG | トーゴ        |
| BG | ブルガリア        | CW | ギニアビサウ      | MK | マケドニア旧ユーゴスラヴィア共和国 | TJ | タジキスタン     |
| BJ | ベナン          | GR | ギリシャ        | ML | マリ                | TM | トルクメニスタン   |
| BR | ブラジル         | HU | ハンガリー       | MN | モンゴル              | TR | トルコ        |
| BY | ベラルーシ        | ID | インドネシア      | MR | モリタニア             | TT | トリニダード・トバゴ |
| CA | カナダ          | IE | アイルランド      | MW | モラウイ              | UG | ウガンダ       |
| CC | 中央アフリカ共和国    | IL | イスラエル       | MX | メキシコ              | US | 米国         |
| CF | コンゴ          | IS | アイスランド      | NE | ニジェール             | UZ | ウズベキスタン    |
| CH | スイス          | IT | イタリア        | NL | オランダ              | VN | ウイエトナム     |
| CI | コート・ジボアール    | JP | 日本          | NO | ノルウェー             | YU | ユーゴスラビア    |
| CM | カメルーン        | KE | ケニア         | NZ | ニュージーランド          | ZW | ジンバブエ      |
| CN | 中国           | KG | キルギスタン      | PL | ポーランド             |    |            |
| CU | キューバ         | KP | 朝鮮民主主義人民共和国 | PT | ポルトガル             |    |            |
| CZ | チェコ共和国       | KR | 大韓民国        | RO | ルーマニア             |    |            |
| DE | ドイツ          | KZ | カザフスタン      | RU | ロシア連邦             |    |            |
| DK | デンマーク        | LC | セントルシア      | SD | スーダン              |    |            |
| EE | エストニア        | LI | リヒテンシュタイン   |    |                   |    |            |